Important Formulas for DRS 113

Chapter 4

Formula for classical probability: $P(E) = \frac{number\ of\ outcomes\ in\ E}{total\ number\ of\ outcomes\ in\ sample\ space} = \frac{n(E)}{n(S)}$

Formula for empirical probability: $P(E) = \frac{frequency\ for\ class}{total\ frequencies\ in\ distribution} = \frac{f}{n}$

Addition rule 1, for two mutually exclusive events: $P(A\ or\ B) = P(A) + P(B)$

Addition rule 2, for events that are not mutually exclusive: $P(A\ or\ B) = P(A) + P(B) - P(A\ and\ B)$

Multiplication rule 1, for independent events: $P(A\ and\ B) = P(A)\,.\,P(B)$

Multiplication rule 2, for dependent events: $P(A\ and\ B) = P(A).\,P(B\backslash A)$

Formula for conditional probability: $P(B\backslash A) = \frac{P(A\ and\ B)}{P(A)}$

Formula for complementary events: $P(\bar{E}) = 1 - P(E)$ or $P(E) = 1 - P(\bar{E})$ or $P(E) + P(\bar{E}) = 1$

Fundamental counting rule: In a sequence of *n* events in which the first one has k_1 possibilities, the second event has k_2 possibilities, the third has k_3 possibilities, etc., the total number of possibilities of the sequence will be

$K_1 . k_2 . k_3 . . . k_n$

Permutation rule 1: The number of permutations of *n* objects taking r objects at a time when order is important is ${}_nP_r = \frac{n!}{(n-r)!}$

Permutation rule 2: The number of permutations of n objects where r_1 objects are identical, r_2 objects are identical, . . . , r_p objects are identical is $\frac{n!}{r_1!r_1!\,.\,.\,.r_p!}$

Combination rule: The number of combinations of r objects selected from n objects when order is not important is ${}_nC_r = \frac{n!}{(n-r)!r!}$

Chapter 5

Formula for the mean of a probability distribution: $\mu = \Sigma X\,.\,P(X)$

Formulas for the variance and standard deviation of a probability distribution:

$$\sigma^2 = \Sigma[X^2.\,P(X)] - \mu^2$$

$$\sigma = \sqrt{\Sigma[X^2.\,P(X)] - \mu^2}$$

Formula for expected value: $E(X) = \Sigma X\,.\,P(X)$

Binomial probability formula: $P(X) = \frac{n!}{(n-X)!X!}\,.\,p^x.\,q^{n-X}$ where X = 0, 1, 2, 3, . . . , n

Formula for the mean of the binomial distribution: $\mu = n\,.\,p$

Formulas for the variance and standard deviation of the binomial distribution:

$$\sigma^2 = n \,.\, p \,.\, q$$

$$\sigma = \sqrt{n \,.\, p \,.\, q}$$

Formula for the multinomial distribution:

$$P(X) = \frac{n!}{X_1!\,.\,X_2!\,.\,X_3!\,\ldots X_k!}\,.\,{P_1}^{X_1}\,.\,{P_2}^{X_2}\ldots{P_k}^{X_k}$$

(The **X**'s sum to ***n*** and the ***p***'s sum to 1.)

Formula for the Poisson distribution: $P(X;\lambda) = \frac{e^{-\lambda}\lambda^X}{X!}$ Where X = 0, 1, 2, . . .

Formula for the hypergeometric distribution: $P(X) = ({}_aC_X \cdot {}_bC_n - X)/({}_{a+b}C_n)$

Formula for the geometric distribution: $P(n) = p(1-p)^{n-1}$ where n = 1, 2, 3, . . .

Chapter 6

Formula for the *z* score (or standard score): $z = \frac{X-\mu}{\sigma}$

Formula for finding a specific data value: $X = z \,.\, \sigma + \mu$

Formula for the mean of the sample means: $\mu_{\bar{X}} = \mu$

Formula for the standard error of the mean: $\sigma_{\bar{X}} = \frac{\sigma}{\sqrt{n}}$

Formula for the *z* value for the central limit theorem: $z = \frac{\bar{X}-\mu}{\sigma/\sqrt{n}}$

Formulas for the mean and standard deviation for the binomial distribution:

$$\mu = n.p$$

$$\sigma = \sqrt{n.p.q}$$

Yangon University Of Economics
Department of Statistics
PGDRS

PGDRS-113

Probability Theory

Week 7

- Exactly 3 = $P(X=3)$
- More than 3 :$P(X>3) = P(X=4) + P(x=5)+\ldots+ P(X=n)$
- 3 or more : $P(X\geq 3)= P(X=3) + P(X=4)+ \ldots.+P(X=n)$

- At least 3 : $P(X\geq 3)= P(X=3) + P(X=4)+ \ldots.+P(X=n)$

  $=1- P(X< 3)$

- Less than3 :$P(X<3) = P(X=0) + P(X=1)+P(X=2)$
- At most 3 : $P(X\leq 3)=P(X=0) +P(X=1)+P(X=2)+ P(X=3)$
  - $= 1- P(X>3)$

A binomial experiment is a probability experiment that satisfies the following four requirements

1. There must be a fixed number of trials.
2. Each trial can have only two outcomes or outcomes that can be reduced to two outcomes. These outcomes can be consider as either success or failure
3. The outcomes of each trial must be independent of one other
4. The probability of a success must remain the same for each trial.

The word **success** does not imply that something **good** or **positive** has occurred.

For example, in a probability experiment, we might want to select 10 people and let *S* represent the number of people who were in an automobile accident in the last six months.

In this case, a success would not be a positive or good thing.

EXAMPLE 5–11

Decide whether each experiment is a binomial experiment. If not, state the reason why.

a. Selecting 20 university students and recording their class rank

b. Selecting 20 students from a university and recording their gender

c. Drawing five cards from a deck without replacement and recording whether they are red or black cards

d. Selecting five students from a large school and asking them if they are on the dean’s list

e. Recording the number of children in 50 randomly selected families

SOLUTION

a. No. There are five possible outcomes: freshman, sophomore, junior, senior, and graduate student.

b. Yes. All four requirements are met.

c. No. Since the cards are not replaced, the events are not independent.

d. Yes. All four requirements are met.

e. No. There can be more than two categories for the answers.

The outcomes of a binomial experiment and the corresponding probabilities of these outcomes are called a **binomial distribution.**

In a Binomial experiment, the probability of exactly X successes in trials is

$$P(X= x) = \frac{n!}{x!(n-x)!} p^x q^{n-x} ; x = 0,1,2, \ldots , n$$

$$P(X=x) = \binom{n}{x} P^x q^{n-x} \quad ; x = 0,1,2,\ldots,n$$

Where

$P(S)$ = symbol for probability of success

$P(F)$ =symbol for probability of failure

P =numerical probability of success

q =numerical probability of failure

n =number of trials

x =numbers of trials in success

EXAMPLE 5–12 Survey on Doctor Visits

A survey found that one out of five Americans says he or she has visited a doctor in any way given month. If 10 people are selected at random, find the probability that exactly 3will have visited a doctor last month?

SOLUTION

In this case, n = 10, X = 3, p = 1/5 , and q = 4/5. Hence,

$$P(X=x) = \binom{10}{x} (1/5)^{x} (4/5)^{n-x} \qquad ; x = 0,1,2,\ldots,10$$

$$P(X= 3) = \binom{10}{3} \left(\frac{1}{5}\right)^{3} \left(\frac{4}{5}\right)^{7} = 0.201$$

So, there is a 0.201 probability that in a random sample of 10 people, exactly 3 of them visited a doctor in the last month.

X!= x * (x-1) * (x-2) …

2! = 2 * 1 =

3!= 3*2*1 = 6

4!= 4*3*2*1 = 24

10C3 * (1/5)^3 *(4/5)^7 = 0.2013

EXAMPLE 5–13 Survey on Employment

A survey from Teenage Research Unlimited (Northbrook, Illinois) found that 30%of teenage consumers receive their spending money from part-time jobs. If 5 teenager are selected at random, find the probability that at least 3 of them will have part time jobs.

Let X be the number of spending money from part-time jobs.

p be the probability of spending money from part-time jobs.

n= number of trial= 5

SOLUTION

To find the probability that at least 3 have part-time jobs, it is necessary to find the individual probabilities for 3, or 4, or 5 and then add them to get the total probability.

$P(X = x) = \binom{5}{x}(0.3)^x(0.7)^{5-x}$; x= 0,1,…,5

$$P(3) = \frac{5!}{(5-3)!3!}(0.3)^3(0.7)^2 \approx 0.132$$

$$P(4) = \frac{5!}{(5-4)!4!}(0.3)^4(0.7)^1 \approx 0.028$$

$$P(5) = \frac{5!}{(5-5)!5!}(0.3)^5(0.7)^0 \approx 0.002$$

Hence,

P(at least three teenagers have part-time jobs)

$= 0.132 + 0.028 + 0.002 = 0.162$

Mean, Variance, and Standard Deviation for the Binomial Distribution

Mean

$$\mu = n.p$$

Variance

$$\sigma^2 = n.p.q$$

Standard Deviation

$$\sigma = \sqrt{n.p.q}$$

EXAMPLE 5–22 Tossing a Coin

A coin is tossed 4 times. Find the mean, variance, and standard deviation of the number of heads that will be obtained.

SOLUTION

With the formulas for the binomial distribution and $n = 4$, $p = 1/2$, and $q = 1/2$, the results are

Mean $\mu = n.p = 4.1/2 = 2$

Variance $\sigma^2 = n.p.q = 4.(1/2).(1/2) = 1$

Standard Deviation $\sigma = \sqrt{n.p.q} = \sqrt{1} = 1$

In this case, the mean is two heads. The variance is 1 and the standard deviation is 1.

EXAMPLE 5–23 Rolling a Die

An 8-sided die (with the numbers 1 through 8 on the faces) is rolled 560 times. Find the mean, variance, and standard deviation of the number of 7s that will be rolled.

SOLUTION

This is a binomial experiment with $n = 560$, $p = 1/8$, and $q = 7/8$ so that

Mean

$$\mu = n.p = 560.1/8 = 70$$

Variance $\sigma^2 = n.p.q = 560.1/8.7/8 = 61.25$

Standard Deviation $\sigma = \sqrt{n.p.q} = \sqrt{61.25} = 7.826$

In this case, the mean of the number of 7s obtained is 70. The variance is 61.25, and the standard deviation is 7.826.

Exercise

1.**Prison Inmates :**Forty percent of prison inmates were unemployed when they entered prison. If 5 inmates are randomly selected, find these probabilities:

a. Exactly 3 were unemployed. *b*. At most 4 were unemployed.

c. At least 3 were unemployed. d. Fewer than 2 were unemployed.

2.Airline Accidents Twenty-five percent of commercial airline accidents are caused by bad weather. If 300 commercial accidents are randomly selected, find the mean, variance, and standard deviation of the number of accidents caused by bad weather.


GOOD NIGHT...!
noexit4u.blogspot.com

Yangon University Of Economics

PGDRS

Department Of Statistics

Probability Theory
PGDRS 113

Presented By

Daw Aye Aye Htwe

Associated Professor

Department Of Statistics

- About 12.5% of restaurant bills are incorrect. If 200 bills are selected at random, find the probability that at least 22 will contain an error. Is this likely or unlikely to occur?

Solution

- Let x be the number of bill is incorrect.

P = the probability of bill is incorrect.

P = 12.5% =0.125 , n = 200

n*p = 0.125*200 = 25 > 5, σ^2 = npq = 200*0.125*0.875 = 21.875, σ = 4.68

X ~ Normally distributed (μ, σ^2 *or* σ)

μ = 25, σ = 4.68

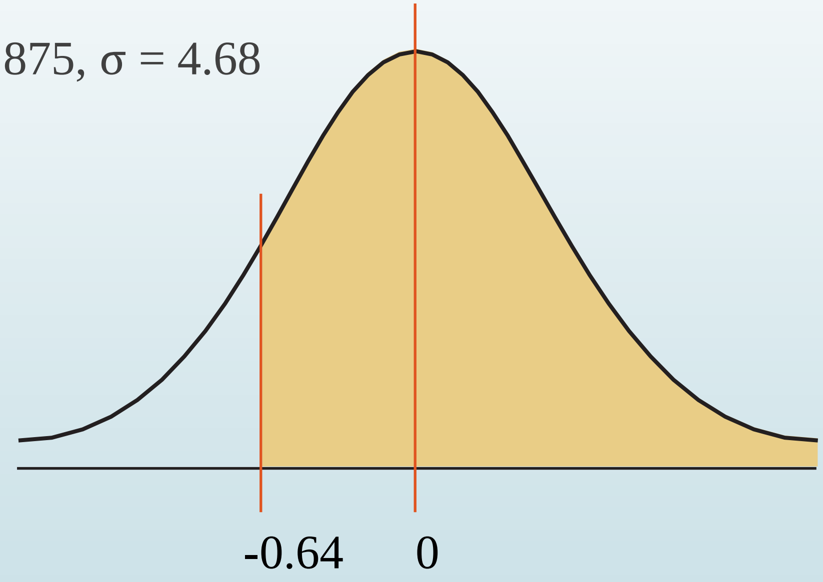
The probability that at least 22 will contain an error

$$P(x \geq 22) = P\left(\frac{x-\mu}{\sigma} \geq \frac{22-\mu}{\sigma}\right) = P\left(z \geq \frac{22-25}{4.68}\right)$$

$= P(z \geq -0.64)$

$= 1 - P(z < -0.64)$

$= 1 - 0.26109 = 0.7389$, This is likely to occur.



- Page- 339, No.26
- A mandatory competency test for high school sophomores has a normal distribution with a mean of 400 and a standard deviation of 100.

(a). The top 3% of students receive $500. What is the minimum score you would need to receive this award?

(b). The bottom 1.5% of students must go to summer school. What is the minimum score you would need to stay out of this group?

- Let x be the number of a mandatory competency test for high school sophomores
- X ~ Normally distributed (μ, σ^2 or σ)
- μ = 400, σ = 100
- (a) x = ? μ= 400 σ = 100

(a)

- $\mu = 400$, $\sigma = 100$

$P(z > A) = 1 - P(z < A)$

$= 1 - 3\%$

$P(z > 1.88) = 0.97$

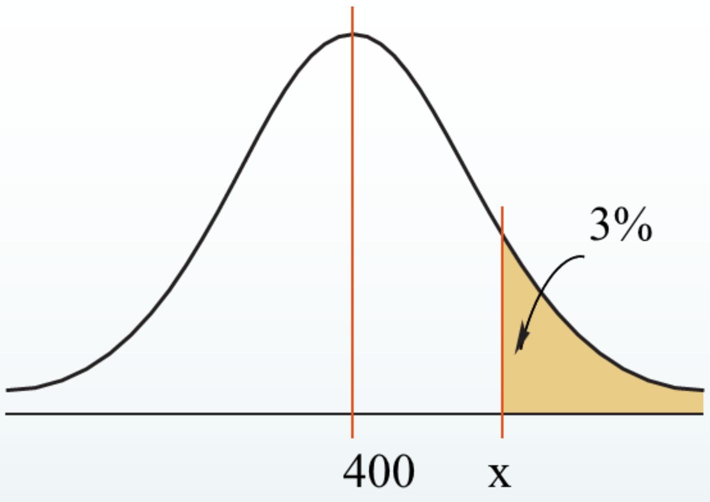
$z = 1.88$

$X = z \cdot \sigma + \mu$

$= (1.88)(100) + 400$

$= 588$

- The minimum score is 588 to receive this award.



- $(b)\ \mu = 400 \qquad \sigma = 100$
- $P(z < A) = 1.5\%$ or 0.015
- $P(z > -2.17) = 0.015$

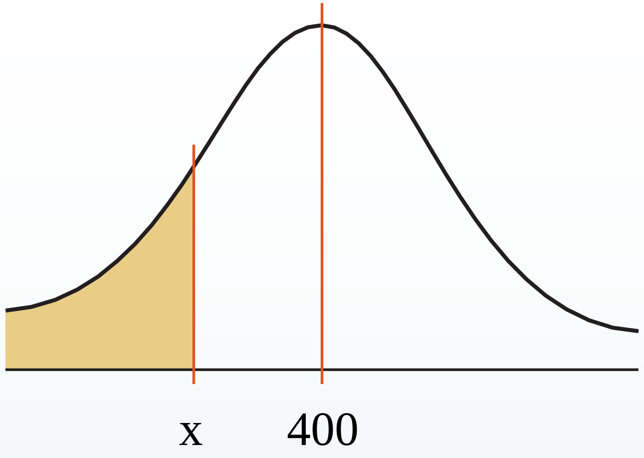
$$z = -2.17$$

$$X = z \cdot \sigma + \mu$$

$$= (-2.17)(100) + 400$$

$$= 183$$

- The minimum score is 183 to stay out of this group



- Page 308, N0.31
- A youth group has 8 boys and 6 girls. If a slate of 4 officers is selected, find the probability that exactly a. 3 are girls b. 2 are girls c. 4 are boys
- Let X be the number of a slate officer.

N = a +b = 6 + 8 = 14, n = 4, a = 6(girl), b = 8 (boy)

X~ Hypergeometric Distribution

$$P(X) = \frac{{}_aC_x \cdot {}_bC_{n-x}}{{}_{a+b}C_n}$$

(a) the probability that exactly 3 are girl = P (X=3) = $\frac{{}_6C_3 \cdot {}_8C_1}{{}_{14}C_4} = \frac{20*8}{1001}$ = 0.1598

(b)the probability that exactly 2 are girls = P(X=2) = $\frac{{}_6C_2 \cdot {}_8C_2}{{}_{14}C_4}$ = 0.4195

© find the probability that exactly 4 are boys = P(X=0) = $\frac{{}_6C_0 \cdot {}_8C_4}{{}_{14}C_4} = \frac{1*70}{1001}$ = 0.0699

- Page 308, N0.28
- If 8% of the population of trees are elm trees, find the probability that in a sample of 100 trees, there are exactly 6 elm trees. Assume the distribution is approximately Poisson.
- Let X be the number of elm tree.

P be the probability of the trees are elm tree , P = 0.08

n=100, n*p = 0.08 *100 = 8 > 5 (Binomial Distribution)

- X~ approximately Poisson distribution

λ be the average of elm tree = 8

- P(X=x; λ) = $\frac{e^{-\lambda}\lambda^{x}}{x!}$, X= 0,1,2,…

$= \frac{e^{-8}(8^{x})}{x!}$, , X= 0,1,2,…..,100

$$P(x=6) = \frac{e^{-8}(8^{6})}{6!} = 0.1221$$

x	P(X=x)	P(X<=x)
0	0.000335	0.0003
1	0.002684	0.0030
2	0.010735	0.0138
3	0.028626	0.0424
4	0.057252	0.0996
5	0.091604	0.1912
6	0.122138	0.3134
7	0.139587	0.4530
8	0.139587	0.5925
9	0.124077	0.7166
10	0.099262	0.8159

Page – 298,(No.3)

According to the manufacturer, M&M's are produced and distributed in the following proportions: 13% brown, 13% red, 14% yellow, 16% green, 20% orange, and 24% blue. In a random sample of 12 M&M's, what is the probability of having 2 of each color?

X~ Multinomial Distributio

$$P(X) = \frac{n!}{X_1!.X_2!.X_3!...X_k!}.P_1^{x_1}.P_2^{x_2}...P_k^{x_k}$$

$X_1 = 2$, $x_2 = 2$, $x_3 = 2$, $x_4 = 2$, $x_5 = 2$, $x_6 = 2$, $n=12$

$P_1=0.13$, $p_2= 0.13$, $p_3= 0.14$, $p_4=0.16$, $p_5= 0.20$, $p_6 = 0.24$

$$P(x) = \frac{12!}{2!2!2!2!2!2!} . (0.13)^2*(0.13)^2*(0.14)^2*(0.16)^2*(0.20)^2*(0.24)^2$$

$$= 0.00247$$

- **Page** -307(No.22)
- Employed Women If 60% of all women are employed outside the home, find the probability that in a sample of 20 women,

a. Exactly 15 are employed

b. At least 10 are employed

c. At most 5 are not employed outside the home

Solution

Let x be the number of a woman is employed outside the home

p be the probability of a woman is employed outside the home

P= 0.60, n = 20

X~ Binomial Distribution

P(X= x) = $\frac{n!}{x!(n-x)!} p^x q^{n-x}$; $x = 0,1,2, \dots, n$

P(X=x) = $\binom{n}{x} P^x q^{n-x}$; x = 0,1,2,…,n

P(X=x) = $\binom{20}{x} (0.6)^x (0.4)^{n-x}$; x = 0,1,2,…,20

Yangon University Of Economics

PGDRS

Department Of Statistics

*Probability Theory PGDRS 113

Presented By

Daw Aye Aye Htwe

Associated Professor

Department Of Statistics

Chapter 4

Probability & Counting Rules

*Introduction

* "The only two sure things in life are death and taxes."

* Decisions are made constantly that are governed at least in part by chance

* **Probability**
  * General concept defined as the chance of an event occurring
  * Used in fields of insurance, investments, gambling, weather forecasting, and various other fields

*4.1 – Sample Spaces & Probability

Processes such as flipping a coin, rolling a die, or drawing a card from a deck are called probability experiments.

- **Probability experiment**
  - Chance process that leads to a well-defined results called outcomes
- **Outcome**
  - Result of a single trial of a probability experiment

*Sample Space

* **Sample space**
  * The set of all possible outcomes of a probability experiment

Sample spaces for various experiments:

Experiment	Sample Space
Toss one coin	Head, Tail
Roll a die	1, 2, 3, 4, 5, 6
Answer to true/false question	True, False
Toss two coins	Head-head, tail-tail, head-tail, tail-head

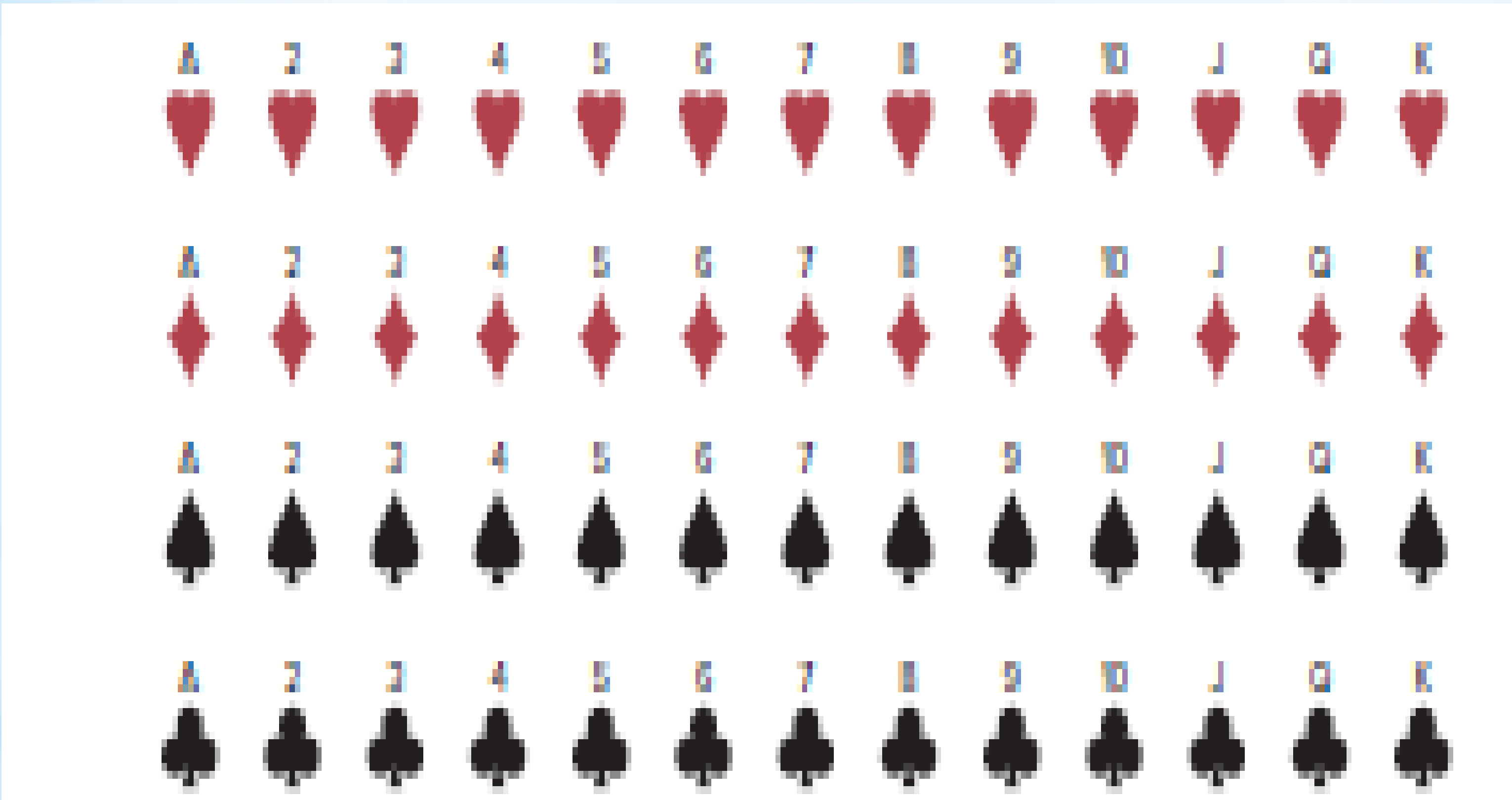
*Examples

* 4-1
  * Find the sample space for rolling two dice
* 4-2
  * Find the sample space for drawing one card from an ordinary deck of cards
* 4-3
  * Find the sample space for the gender of the children if a family has three children

4-1 Rolling two dice

			Die 2			
Die 1	1	2	3	4	5	6
1	(1,1)	(1,2)	(1,3)	(1,4)	(1,5)	(1,6)
2	(2,1)	(2,2)	(2,3)	(2,4)	(2,5)	(2, 6)
3	(3,1)	(3,2)	(3,3)	(3, 4)	(3, 5)	(3, 6)
4	(4,1)	(4,2)	(4, 3)	(4, 4)	(4, 5)	(4, 6)
5	(5,1)	(5,2)	(5, 3)	(5, 4)	(5, 5)	(5, 6)
6	(6,1)	(6,2)	(6, 3)	(6, 4)	(6, 5)	(6, 6)

sample space = S = {(1,1), (1,2), (1,3), … , (6,6)}


A 2 3 4 5 6 7 8 9 10 J Q K
A 2 3 4 5 6 7 8 9 10 J Q K
A 2 3 4 5 6 7 8 9 10 J Q K
A 2 3 4 5 6 7 8 9 10 J Q K

ck of
cards

4-3 Gender of the children if a family has three children

There are two genders, boy and girl, and each child could be either gender. Hence, there are eight possibilities, as shown here.

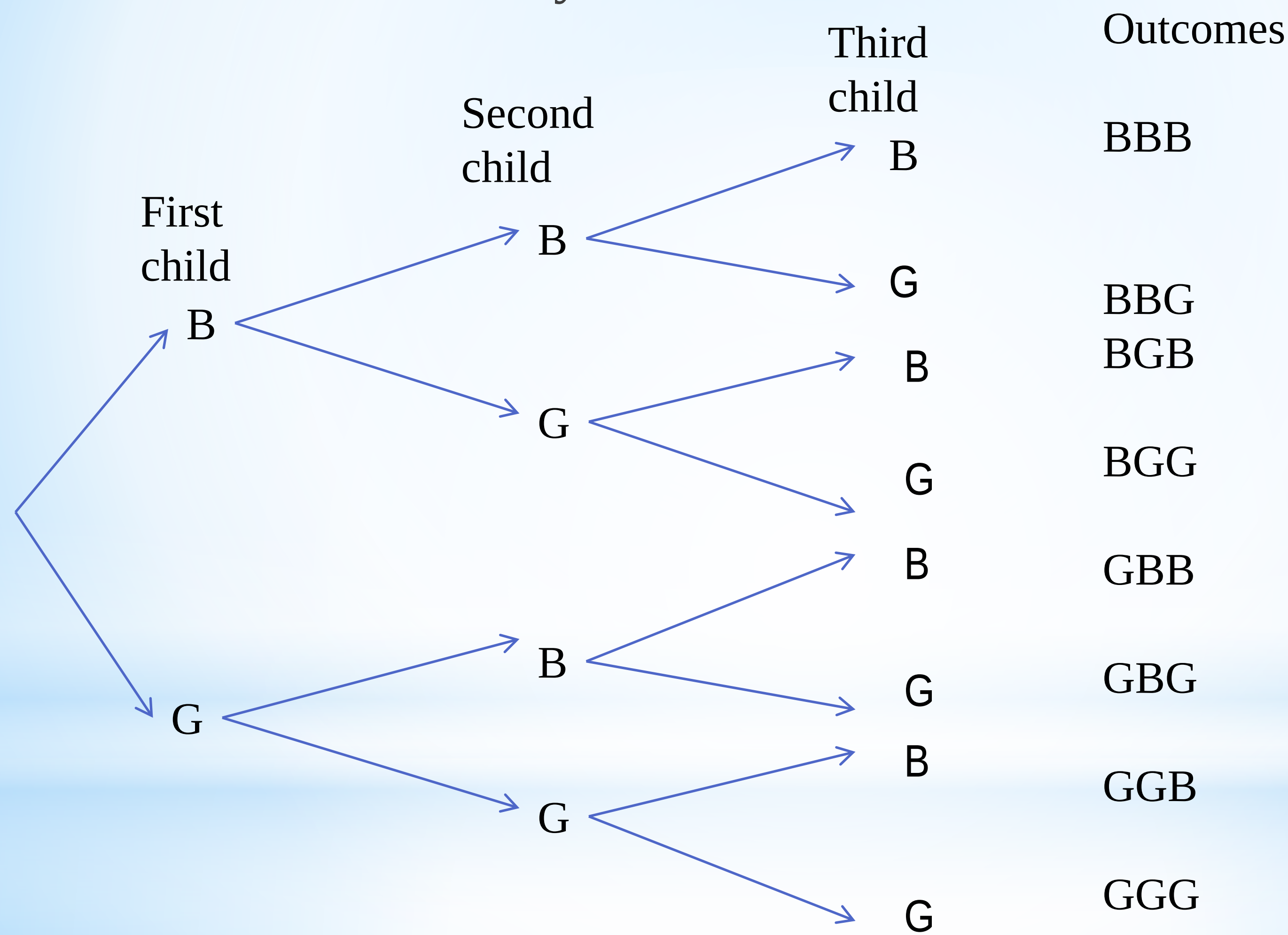
BBB BBG BGB GBB GGG GGB GBG BGG

sample space = S ={(BBB), (BBG), (BGB), (GBB), (GGG), (GGB), (GBG), (BGG)}

Tree Diagram

* **Tree diagram**
  * Device consisting of line segments emanating from a starting point and also from the outcome point, used to determine all possible outcomes of an event
* Example 4-4
  * Make a tree diagram for gender of children in a family with three children

Example 4-4 Tree diagram for gender of children in a family with three children



Types of Probabilities

- **Event**
  - Set of outcomes of a probability experiment
- **Simple event**
  - Event with one outcome (rolling a 6)
- **Compound event**
  - Consists of two or more outcomes or simple events (rolling an even number)

Three Interpretations of Probability

1. Classical Probability
2. Empirical Probability
3. Subjective Probability

*Classical Prob

- Uses sample spaces to determine numer
  probability that an event will happen
- Assumes that all outcomes in sample space are equally likely to happen

Formula for Classical Probability

$$P(E) = \frac{\#\ of\ outcomes\ in\ event\ E}{\#\ of\ outcomes\ in\ sample\ space} = \frac{n(E)}{n(S)}$$

*Expression of Probabilities

* Probabilities can be expressed as decimals, fractions, or percentages

* **Rounding rule**
  * Expressed as a reduced fraction or as a decimal rounded to three decimal places

* 4-5
  * Find probability of getting a black 6

* 4-6

There are 52 cards in an ordinary deck, and there are two black 6s, that is, the 6 of clubs and the 6 of spades. Hence, the probability of getting a black 6 is

$$\approx 0.038$$

$$\frac{2}{52} = \frac{1}{26}$$

4-5 Find probability of getting a black 6

The sample space for the gender of three children has eight outcomes BBB, BBG, BGB, GBB, GGG, GGB, GBG, and BGG. The probability of having exactly two boys is

$$\frac{3}{8}.$$

- **4-6 If a family has three children, find probability that two of the three children are girls**

*“And” & “Or” in Probability

* It is important to understand the meaning of the words “and” and “or” in probability theory
* “and” means at the same time (multiply)
* “or” could be inclusive or exclusive depending on the circumstances of the event (addition)

* A card is drawn from an ordinary deck, find these probabilities

a) Getting a heart

b) Getting a black card

c) Getting the 8 of diamonds

d) Getting a queen

e) Getting a face card

*Example 4-7

a. There are 13 hearts in a deck of 52 cards; hence,

$P(\text{heart}) = \frac{13}{52} = 0.25$

b. There are 26 black cards in a deck, that is, 13 clubs and 13 spades. So the probability is

$P(\text{black card}) = \frac{26}{52} = 0.5$

c. There is one 8 of diamonds in a deck of 52 cards, so the probability is

$P(\text{8 of diamonds}) = \frac{1}{52} \approx 0.019$

d. There are four queens in a deck of 52 cards; hence,

$P(\text{queen}) = \frac{4}{52} \approx 0.077$

e. There are 12 face cards in an ordinary deck of cards, that is, 4 suits (diamonds, hearts, spades, and clubs) and 3 face cards of each suit (jack, queen, and king), so

$P(\text{face card}) = \frac{12}{52} \approx 0.231$

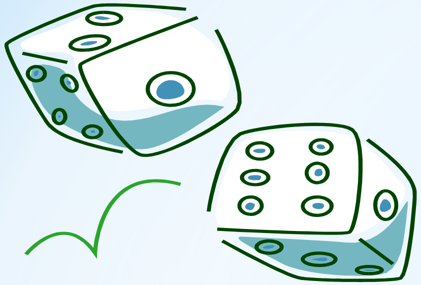
*Probability Rules

1. The probability of an event E is a number between and including 0 and 1, denoted by $0 \leq P(E) \leq 1$

2. If an event E cannot occur, its probability is 0

3. If an event E is certain, then the probability of E is 1

4. The sum of the probabilities of all the outcomes in the sample space is 1



*Examples

* 4-8
  * When a single die is rolled, find the probability of getting a 9

P(9) = $\frac{0}{6} = 0$

* 4-9
  * When a single die is rolled, what is the probability of getting a number less than 7

P(number less than 7) = $\frac{6}{6} = 1$

* The complement of an event E is the set of outcomes in the sample space that are not included in the outcomes of event E. The complement of E is denoted by $\bar{E}$ (read E bar).

* Complementary Events

Example 4-10

Find the complement of each event:

a) Selecting a month that has 30 days.

b) Selecting a day of the week that begins with the letter S.

c) Rolling two dice and getting a number whose sum is 7.

d) Selecting a letter of the alphabet(excluding y) that is a vowel.

Solution

a). Selecting a month that has 28 or 31 days, that is, January, February, March, May, July, August, October, or December.

b). Selecting a day of the week that does not begin with S, that is, Monday, Tuesday, Wednesday, Thursday, or Friday.

c). Rolling two dice and getting a number whose sum is not 7, that is, 2,3,4,5,6,8,9,10,11, or 12.

d). Selecting a letter of the alphabet that is a consonant.

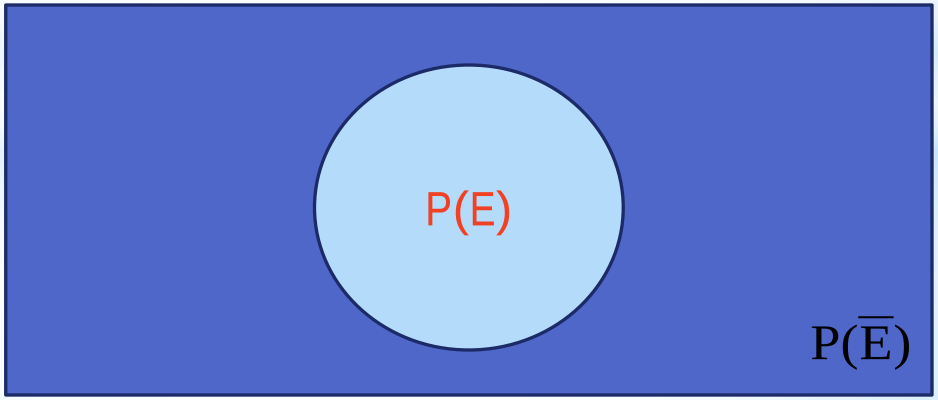
*Rules for Complementary Events

$$P(\overline{E}) = 1 - P(E)$$

$$P(E) = 1 - P(\overline{E})$$

$$P(E) + P(\overline{E}) = 1$$

* If the probability of an event or the probability of its complement is known, then the other can be found by subtracting the probability from 1.



- Example 4-11

In a study ,it was found that 24% of people who were victims of a violent crime were age 20 to 24. If a person is selected at random, find the probability that the person is younger than 20 or older than 24.

Solution

Let A = the event that the person is aged 20 to 24

$\overline{A}$ = the event that the person is not aged 20 to 24

$$P(A) = 24\% = 0.24$$

The probability that the person is younger than 20 or older than 24.

$$P(\overline{A}) = 1 - P(A)$$

$$P(\overline{A}) = 1 - 0.24 = 0.76 = 76\%$$

Empirical Probability

* Empirical probability
  * Relies on actual experience to determine the likelihood of outcomes

Formula for Empirical Probability

* Given a frequency distribution, the probability of an event being in a given class is

$$P(E) = \frac{frequency\ for\ the\ class}{total\ frequencies\ in\ the\ distribution} = \frac{f}{n}$$

* Empirical probabilities are based on observation

*Empirical Probabilities Examples

* 4-12
  * In the following travel survey, find the probability that a person will travel by train or bus over the Thanksgiving holiday.

Method	Frequency
Drive	41
Fly	6
Train or Bus	3

Solution

The probability that a person will travel by train or bus over the Thanksgiving holiday.

$$P(E) = \frac{f}{n} = \frac{3}{50} = 0.06$$

In a sample of 50 people, 21 had type O blood, 22 had Type A, 5 had type B, and 2 had type AB. Set up a frequency distribution and find the following probabilities:

a) A person has type O blood
b) A person has type A or type B blood
c) A person has neither type A nor type O blood
d) A person does not have type AB blood

Solution

Frequency Distribution

Type	Frequency
A	22
B	5
AB	2
O	21
Total	50

a. $$P(O) = \frac{f}{n} = \frac{21}{50}$$

b. $$P(A \text{ or } B) = \frac{22}{50} + \frac{5}{50} = \frac{27}{50}$$

c. $$P(\text{ neither A nor O}) = \frac{5}{50} + \frac{2}{50} = \frac{7}{50}$$

d. $$P(\text{not AB}) = 1 - P(AB) = 1 - \frac{2}{50} = \frac{48}{50}$$


THANK YOU
For Your Attention!

YANGON UNIVERSITY OF ECONOMICS

DEPARTMENT OF STATISTICS

PGDRS

PROBABILITY THEORY
PGDRS 113

Presented By

Daw Aye Aye Htwe

Associated Professor

Department Of Statistics

Chapter 4

Probability & Counting Rules

Probability ? **Statistics**

Science of chance, uncertainties

what is possible ,
what is probable

Science of data

collecting, processing,
presentation, analysing
interpretation of data

mathematical formulas **numbers with context**

EMPIRICAL PROBABILITY

- **Empirical probability**
  - Relies on actual experience to determine the likelihood of outcomes

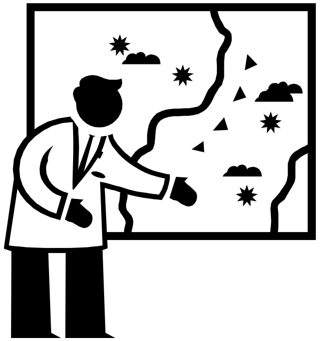
<u>Formula for Empirical Probability</u>

  - Given a frequency distribution, the probability of an event being in a given class is

$$P(E) = \frac{frequency\ for\ the\ class}{total\ frequencies\ in\ the\ distribution} = \frac{f}{n}$$

  - Empirical probabilities are based on observation

SUBJECTIVE PROBABILITY



Subjective probability

- Uses a probability value based on an educated guess or estimate

Guesses are based on a person's experience and evaluation of a solution

All three types of probability (classical, empirical, subjective) are used to solve a variety of problems in business, engineering, and other fields

4.2 – ADDITION RULES FOR PROBABILITY

Many problems involve finding probability of two or more events

For example, probability person is a female or a Republican has 3 possibilities

1. Person is a female
2. Person is a Republican
3. Person is both female and Republican

Another example, probability person is a Democrat or Independent, only 2 possibilities

1. Person is a Democrat
2. Person is an Independent

MUTUALLY EXCLUSIVE

Female or Republican example is not *mutually exclusive*

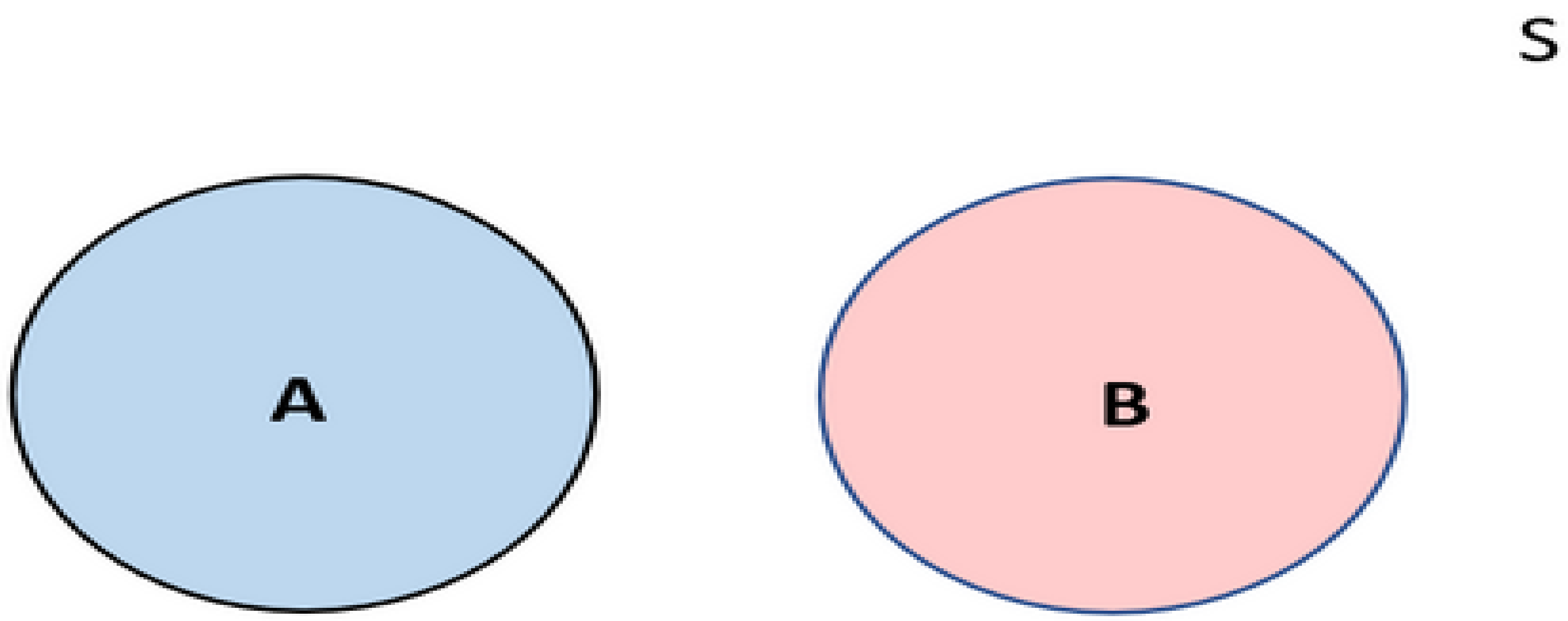
Democrat or Independent example is *mutually exclusive*

Mutually exclusive events

- Events that cannot occur at same time

- Have no outcomes in common

Mutually Exclusive Events



Example 4-15

Determine whether the two events are mutually exclusive. Explain your answer.

a. Randomly selecting a female student

Randomly selecting a student who is a junior

b. Randomly selecting a person with type A blood

Randomly selecting a person with type O blood

c. Rolling a die and getting an odd number

Rolling a die and getting a number less than 3

d. Randomly selecting a person who is under 21 years of age

Randomly selecting a person who is over 30 years of age

SOLUTION

a. These events are not mutually exclusive since a student can be both female and a junior.

b. These events are mutually exclusive since a person cannot have type A blood and type O blood at the same time.

c. These events are not mutually exclusive since the number 1 is both an odd number and a number less than 3.

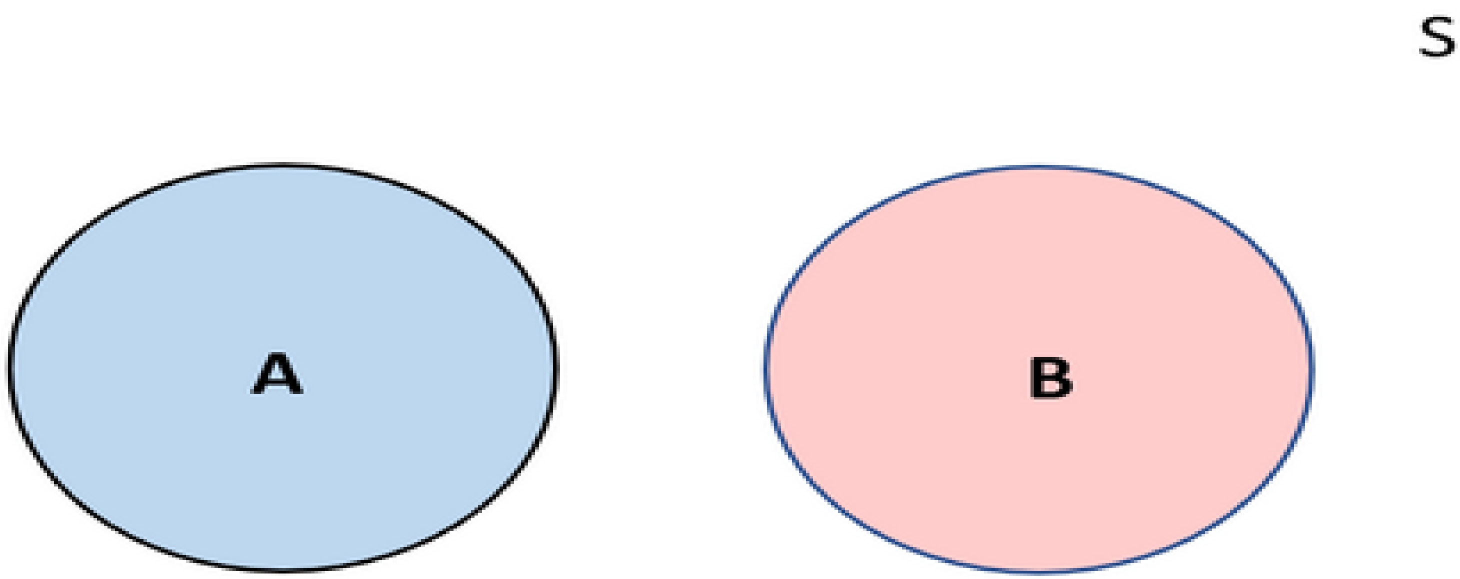
d. These events are mutually exclusive since a person cannot be both under 21 and over 30 years of age at the same time.

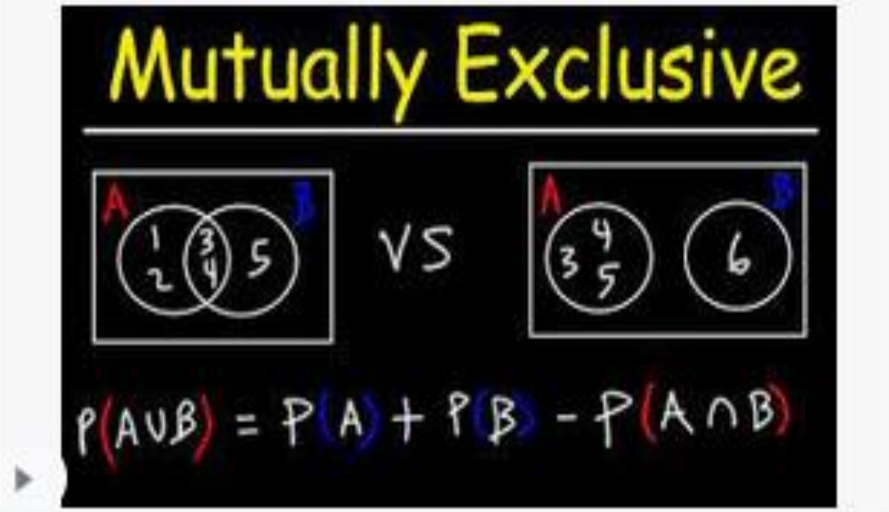
ADDITION RULE 1

- When two events A and B are mutually exclusive, probability that A or B will occur is

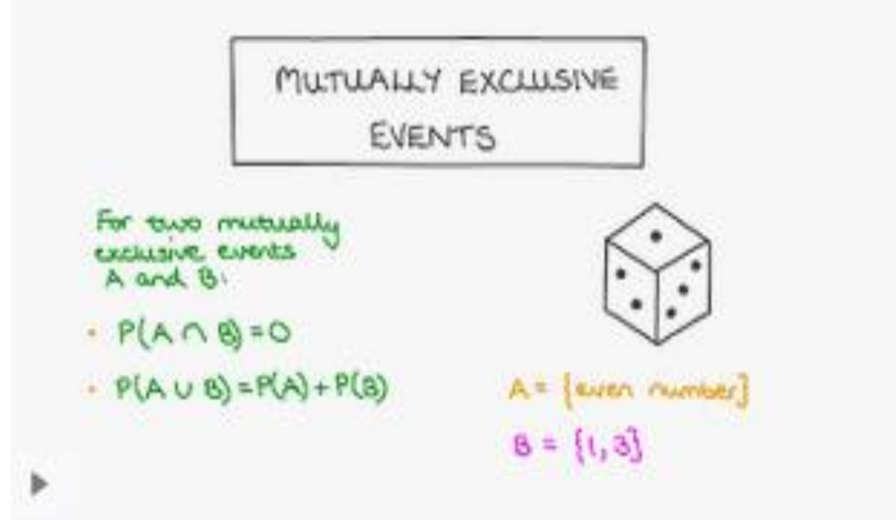
$$P(A \text{ or } B) = P(A) + P(B)$$

Mutually Exclusive Events

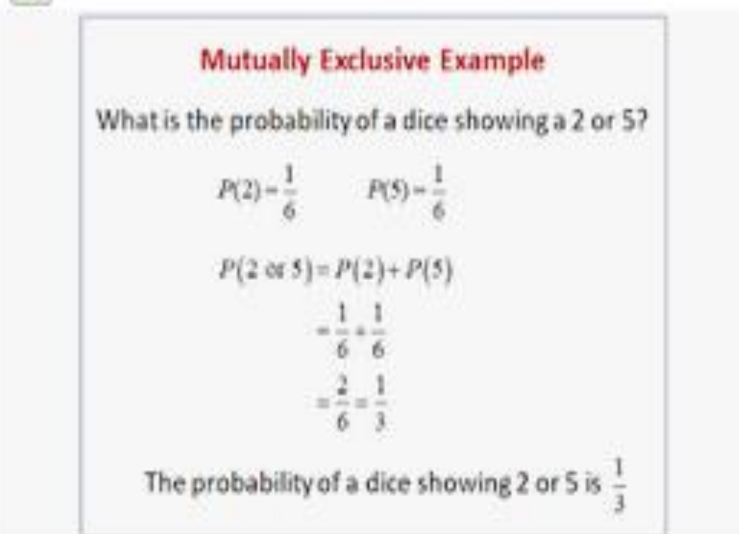




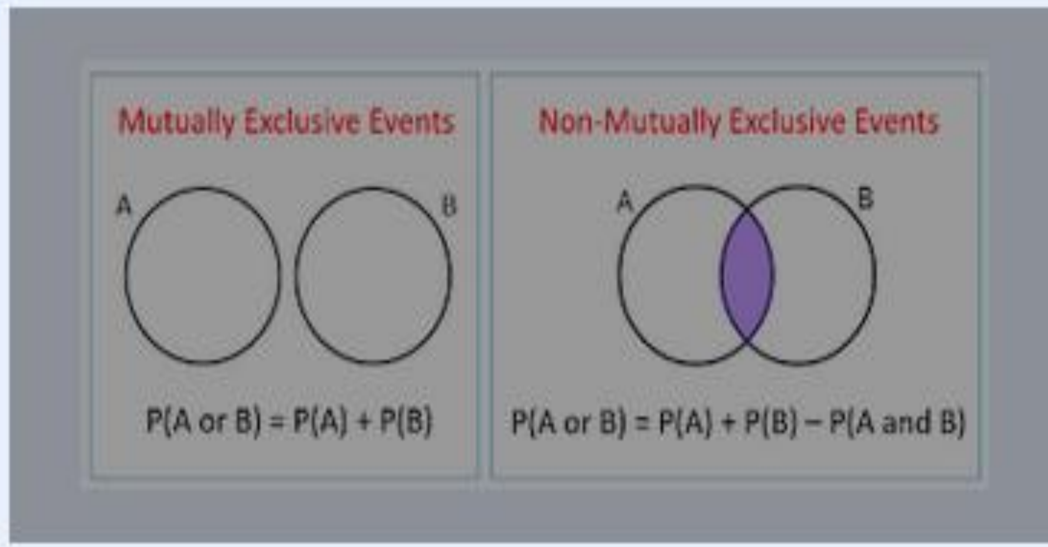
Probability of Mutually Exclusive ...
youtube.com



Lesson Video: Mutually Exclusive Events ...
nagwa.com



Mutually Exclusive Events (video ...
onlinemathlearning.com



Mutually Exclusive Events (video ...
onlinemathlearning.com

- **Example 4-17**

In the United States there are 59 different species of mammals that are endangered, 75 different species of birds that are endangered, and 68 species of fish that are endangered. If one animal is selected at random, find the probability that it is either a mammal or a fish.

Solution

Since there are 59 species of mammals and 68 species of fish that are endangered and a total of 202 endangered species,

P(mammal or fish) = P(mammal) + P(fish)

$$= \frac{59}{202} + \frac{68}{202} = \frac{127}{202}$$

$$= 0.629$$

The events are mutually exclusive.

Example 4-18

The corporate research and development center for three local companies have the following numbers of employees:

U.S. Steel	**110**
Alcoa	**750**
Bayer Material Science	**250**

If a research employee is selected at random, find the probability that the employee is employed by U.S. Steel or Alcoa.

Solution

P(U.S. Steel or Alcoa) = P(U.S. Steel) + P(Alcoa)

$$= \frac{110}{1110} + \frac{750}{1110} = \frac{860}{1110} = \mathbf{0.775}$$

- EXAMPLE 4-19

In a survey, 8% of the respondents said that their favourite ice cream flavour is cookies and cream, and 6% like mint chocolate chip. If a person is selected at random, find the probability that her or his favourite ice cream flavour is either cookies and cream or mint chocolate chip.

Solution

$P(\text{cookies and cream or mint chocolate chip})$

$= P(\text{cookies and cream}) + P(\text{mint chocolate chip})$

$= 0.08 + 0.06 = 0.14 = 14\%$

These events are mutually exclusive.

ADDITION RULE 2

- If two events A and B are **not** mutually exclusive, then the probability that A or B will occur is

$$\boldsymbol{P(A\ or\ B) = P(A) + P(B) - P(A\ and\ B)}$$

EXAMPLE

4-20

A single card is drawn from an ordinary deck of cards. Find the probability that it is either a 9 or a diamond

Solution

There are 4 nines and 13 diamonds in a deck of cards, and one of the 9s is a diamond, so it is counted twice. Hence,

P(9 or a diamond) = P(9) + P(diamond) −P(9 and a diamond)

$$= \frac{4}{52} + \frac{13}{52} - \frac{1}{52} = \frac{16}{52}$$

$$\approx 0.308$$

EXAMPLE

- **4-21 In a hospital unit there are 8 nurses and 5 physicians; 7 nurses and 3 physicians are females. If a staff person is selected, find the probability that the subject is a nurse or a male.**

Solution

Staff	Females	Males	Total
Nurses	7	1	8
Physicians	3	2	5
Total	10	3	13

The probability is

$$P(\textit{nurse or male}) = P(\textit{nurse}) + P(\textit{male}) - P(\textit{male nurse})$$

$$= \frac{8}{13} + \frac{3}{13} - \frac{1}{13} = \frac{10}{13} \approx 0.769$$

EXAMPLE

- **4-22**

On New Year's Eve, the probability of a person driving while intoxicated is 0.32, the probability of a person having a driving accident is 0.09, and the probability of a person having a driving accident while intoxicated is 0.06. What is the probability of a person driving while intoxicated or having a driving accident?

Solution

$$\begin{aligned} \boldsymbol{P(intoxicated\ or\ accident)} &= \boldsymbol{P(intoxicated) + P(accident)} \\ &\quad \boldsymbol{-\ P(intoxicated\ and\ accident)} \\ &= \mathbf{0.32 + 0.09 - 0.06 = 0.35} \end{aligned}$$

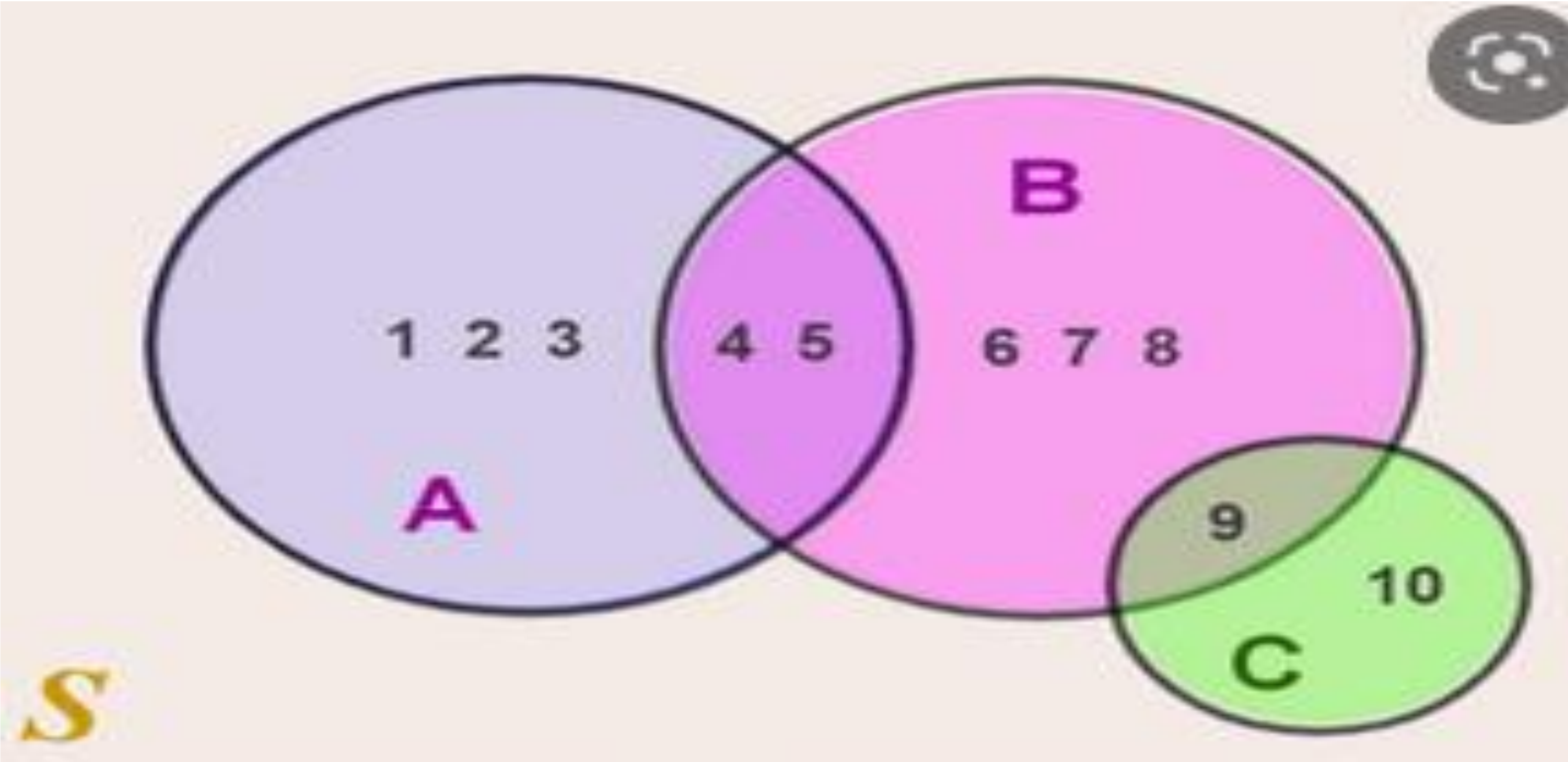
EXTENDING TO 3+ EVENTS

- The probability rules can be extended to three or more events
- For three mutually exclusive events A, B, and C.

$$P(A \text{ or } B \text{ or } C) = P(A) + P(B) + P(C)$$

- For three events that are not mutually exclusive,

$$P(A \text{ or } B \text{ or } C) = P(A) + P(B) + P(C) - P(A \text{ and } B)$$
$$- P(A \text{ and } C) - P(B \text{ and } C) + P(A \text{ and } B \text{ and } C)$$


B
1 2 3
4 5
6 7 8
A
9
10
C
S


Thank You and Good Night
Patrick
McDon

Yangon University Of Economics

Department Of Statistics

PGDRS

Probability Theory
PGDRS 113

Presented By
Daw Aye Aye Htwe
Associated Professor
Department Of Statistics

4.3 – Multiplication Rules & Conditional Probability

- Multiplication rules can be used to find probability of two or more events that occur in a sequence

- Events are either independent or dependent

Independent Events

- **Independent events**
  - Two events A and B are independent if fact that A occurs does not affect probability of B occurring

- Examples of independent events
  - Rolling a die and getting a 6, then rolling another die and getting a 3
  - Drawing a card from a deck and getting a queen, replacing it, and drawing a second card and getting a queen

Multiplication Rule 1

- When two events A and B are **independent**, the probability of both occurring is

$$P(A \text{ and } B) = P(A) * P(B)$$

- Independent events can also be considered as probabilities solved **with replacement**

- 4-23

A coin is flipped and a die is rolled. Find the probability of getting a head on the coin and a 4 on the die.

Solution

The sample space for the coin is H, T; and for the die it is 1, 2, 3, 4, 5, 6.

$$P(\text{head and } 4) = P(\text{head}) \cdot P(4) = \frac{1}{2} \cdot \frac{1}{6} = \frac{1}{12} \approx 0.083$$

Example

- 4-25
  - An urn contains 2 red balls,5 blue balls, and 3 white balls. A ball is selected, its color noted, then it is replaced. A second ball is selected and its color noted. Find the probability of each of these

a. Selecting 3 blue balls
b. Selecting 1 white ball then 1 red ball
c. Selecting 2 blue ball then 1 white ball

Solution

a. $P(\text{blue and blue and blue}) = P(\text{blue}) \cdot P(\text{blue}) \cdot P(\text{blue})$

$$= \frac{5}{10} \cdot \frac{5}{10} \cdot \frac{5}{10} = \frac{125}{1000}$$

$= 0.125$

b. $P(\text{white and red}) = P(\text{white}) \cdot P(\text{red})$

$$= \frac{3}{10} \cdot \frac{2}{10} = \frac{6}{100}$$

$= 0.06$

c. $P(\text{blue and blue and white}) = P(\text{blue}) \cdot P(\text{blue}) \cdot P(\text{white})$

$$= \frac{5}{10} \cdot \frac{5}{10} \cdot \frac{3}{10} = \frac{75}{1000}$$

$= 0.075$

Example

- 4-24
  - A card is drawn from a deck and replaced; then a second card is drawn. Find the probability of getting a king and then a 7.

Solution

$$P(\textit{king and then a 7}) = \frac{4}{52} \cdot \frac{4}{52} = \frac{16}{2704}$$

$$\approx 0.006.$$

Example

- 4-26

  It was found that 3 out of every 4 people who commit a bank robbery are apprehended. If 3 bank robberies are selected at random, find the probability that all three robbers will be apprehended.

Solution

Let *R be the event that a bank robber is apprehended.*

$$P(R \text{ and } R \text{ and } R) = P(R) \cdot P(R) \cdot P(R) = \frac{3}{4} \cdot \frac{3}{4} \cdot \frac{3}{4} = \frac{27}{64}$$

$$\approx 0.422$$

There is about a 42% chance that all 3 robbers will be apprehended.

Dependent Events

- **Dependent events**
  - Two events A and B are **dependent** when the outcome or occurrence of event A affects the outcome or occurrence of event B in such a way that the probability is changed.

- Examples of dependent events
  - Drawing a card from a deck, not replacing it, and drawing a second card
  - Selecting a ball from an urn, not replacing it, and selecting a second ball
  - Having high grades and getting a scholarship
  - Parking in a no-parking zone and getting parking ticket

Conditional Probability

- Probabilities involving dependent events are called **conditional probabilities**

- **Conditional probability**
  - Probability of event B in relationship to an event A is probability that event B occurs given that A has already occurred

Formula for Conditional Probability

- The probability that the second event B occurs given that the first event A has occurred can be found by dividing the probability that both events occurred by the probability that the first event has occurred. The formula is

$$\boldsymbol{P(B|A) = \frac{P(A\ and\ B)}{P(A)}}$$

Multiplication Rule 2

- When two events A and B are dependent, the probability of both occurring is

$$P(A \text{ and } B) = P(A).P(B|A)$$

➢ Dependent events can also be considered as probabilities solved without replacement

Example

- 4-28
  - For a specific year, 5.2% of U.S. workers were unemployed. During that time, 33% of those who were unemployed received unemployment benefits. If a person is selected at random, find the probability that she or he received unemployment benefits if the person is unemployed.

Solution

$$P(\text{unemployed benefits and unemployed}) = P(U) \cdot P(B|U)$$
$$= (0.052)(0.33)$$
$$= 0.017$$

Example

- 4-30
  - Three cards are drawn from a deck and not replaced. Find the probability of these events

a. Getting 3 Jacks

b. Getting an Ace, a King, and a Queen in order

c. Getting a club, a spade, and a heart in order

d. Getting 3 clubs

Solution

- *a. P(3 jacks) =* $\frac{4}{52}\cdot\frac{3}{51}\cdot\frac{2}{50}=\frac{24}{132600}$ ≈ *0.0002*
- *b. P(ace and king and queen) =* $\frac{4}{52}\cdot\frac{4}{51}\cdot\frac{4}{50}=\frac{64}{132600}$ ≈ *0.0005*
- *c. P(club and spade and heart) =* $\frac{13}{52}\cdot\frac{13}{51}\cdot\frac{13}{50}=\frac{2197}{132600}$ ≈ *0.017*
- *d. P(3 clubs) =* $\frac{13}{52}\cdot\frac{12}{51}\cdot\frac{11}{50}=\frac{1716}{132600}$ ≈ *0.013*

Example

- 4-29
  - World Wide Insurance Company found that 53% of the residents of a city had homeowner's insurance (H) with the company. Of these clients, 27% also had automobile insurance (A) with the company. If a resident is selected at random, find the probability that the resident has both homeowner's and automobile insurance with World Wide Insurance

    Company.

Solution

$P(H \text{ and } A) = P(H) \cdot P(A|H) = (0.53)(0.27) = 0.1431 \approx 0.143$

Example

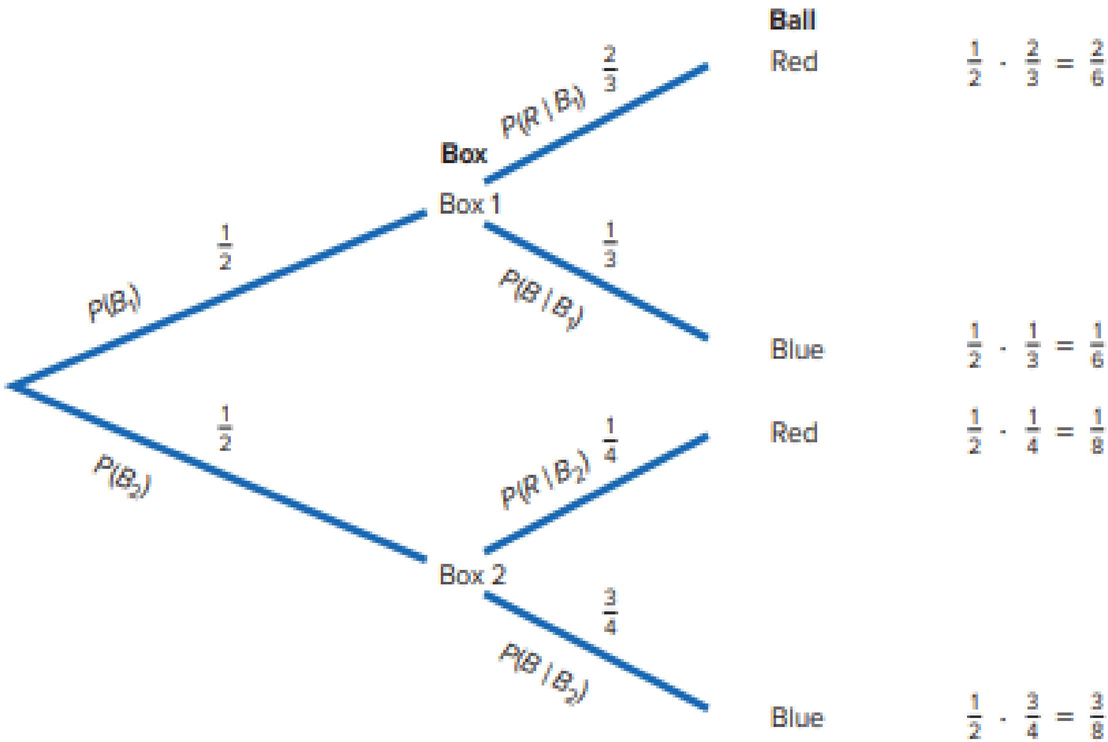
- 4-31

  Box 1 contains 2 red balls and 1 blue ball. Box 2 contains 3 blue balls and 1 red ball. A coin is tossed. If it falls heads up, box 1 is selected and a ball is drawn. If it falls tails up, box 2 is selected and a ball is drawn. Find the probability of selecting a red ball.

Solution

$$\text{P(Red)} = \text{P}(R\,and\,B_1) + P(R\,and\,B_2) = \frac{2}{6} + \frac{1}{8} = \frac{11}{24}$$

FIGURE 4–8 Tree Diagram for Example 4–31



Formula for Conditional Probability

- The probability that the second event B occurs given that the first event A has occurred can be found by dividing the probability that both events occurred by the probability that the first event has occurred. The formula is

$$\boldsymbol{P(B|A) = \frac{P(A\ and\ B)}{P(A)}}$$

If event A and B are dependent

- $P(B|A) = \dfrac{P\ (\ A \text{ and } B)}{P(A)}$

$P(\ A \text{ and } B) = P(A)\ .\ P(\ B|A)$

- $P\ (A|B) = \dfrac{P(\ A \text{ and } B)}{P(B)}$

$P(\ A \text{ and } B) = P(B).\ P(\ A|B)$

Example

- 4-32
  - A box contains black chips and white chips. A person selects two chips without replacement. If the probability of selecting a black chip *and* a white chip is 15/56, and the probability of selecting a black chip on the first draw is 3/8, find the probability of selecting the white chip on the second draw, *given* that the first chip selected was a black chip.

Solution

Let

B = selecting a black chip W = selecting a white chip

$$P(W\,|B) = \frac{P(\,B\text{ and }W)}{P(B)} = \frac{15/56}{3/8}$$

$$\approx 0.714$$

Hence, the probability of selecting a white chip on the second draw given that the first chip selected was black is 0.714.

Example

- 4-33
  - The probability that Sam parks in a no-parking zone *and* gets a parking ticket is 0.06, and the probability that Sam cannot find a legal parking space and has to park in the no-parking zone is 0.20. On Tuesday, Sam arrives at school and has to park in a no-parking zone. Find the probability that he will receive a parking ticket

Solution

Let

N = parking in a no-parking zone T = getting a ticket

$$P(T \mid N) = \frac{P(\text{N and T})}{P(N)} = \frac{0.06}{0.20}$$

$$= 0.30$$

Hence, Sam has a 0.30 probability or 30% chance of getting a parking ticket, given that he parked in a no-parking zone.

Example

- 4-34

  A recent survey asked 100 people if they thought women in the armed forces should be permitted to participate in combat. The results of the survey are shown.

Gender	Yes	No	Total
Male	32	18	50
Female	8	42	50
Total	40	60	100

Find these probabilities.

a. The respondent answered yes, given that the respondent was a female.

b. The respondent was a male, given that the respondent answered no.

Solution

Let

M = respondent was a male, Y = respondent answered yes

F = respondent was a female, N = respondent answered no

$$a.\quad P(Y\,|F) = \frac{\mathrm{P(F\ and\ Y)}}{\mathrm{P(F)}} = \frac{8/100}{50/100}$$

$$= 0.16$$

$$b.\quad P(M\,|N) = \frac{\mathrm{P(N\ and\ M)}}{\mathrm{P(N)}} = \frac{18/100}{60/100}$$

$$= 0.3$$


hank You
AND
Good Nigh

Yangon University of Economics
Department of Statistics
PGDRS.113
Probability Theory

Chapter 4
Probability and Counting Rules
Exercises

Roll two dice and multiple the numbers together.

a. Write out the sample space.

b. What is the probability that the product is a multiple of 6?

c. What is the probability that the product is less than 10?

Solution

a. Sample space = S= { 1, 2, 3, 4, 5, 6,   2, 4, 6, 8,10,12,
3, 6, 9, 12, 15, 18,   4, 8, 12, 16, 20, 24,
5, 10, 15, 20, 25, 30,   6, 12, 18, 24, 30, 36}
= 36

b. P(the product is a multiple of 6) = 4/36

c. P(the product is less than 10) = 17/ 36

Human blood is grouped into four types. The percentage of Americans with each types are listed below.

O =43% A =40% B= 12% AB= 5%

Choose one American at random. Find the probability that this person

(a) Has type B blood

(b) Has type AB or O blood

(c) Does not have type O blood

Solution

(a) P(the person has type B blood) = 0.12

(b) P(the person has type AB or O blood) = 0.05 + 0.43 =0.48

(c) P(the person does not have type O blood) = 1 – P(O)

= 1- 0.43

= 0.57

- At a community swimming pool there are 2 managers, 8 lifeguards, 3 concession stand clerks and 2 maintenance people. If a person is selected at random, find the probability that the person is either a lifeguard or a manager.

Solution

Let

M be the event that the person is a manager.

L be the event that the person is a lifeguard.

C be the event that the person is a clerks.

N be the event that the person is a maintenance people.

Probability distribution of a community swimming pool

Type	frequencies	Probability
Manager (M)	2	0.1333
Lifeguard (L)	8	0.5333
Clerk (C)	3	0.2
Maintenance People (N)	2	0.1333
Total	N= 15	0.9999~ 1

The probability that the person is either a lifeguard or a manager

P (L) = f/n

= P(L or M)

= (8/15 + 2/15)

= (10/ 15)

The probability that the person is neither a lifeguard nor a manager .

= 1- P(L or M) = 1- 10/ 15= 5/15

- A recent study of 200 nurses found that of 125 female nurses, 56 had bachelor’s degrees: and of 75 male nurses, 34 had bachelors degree’s. If a nurse is selected at random, find the probability that the nurse is

(a) A female nurse with a bachelor’s degree

(b) A male nurse

(c) A male nurse with a bachelor’s degree

(d) Based on your answers to parts a, b, and c, explain which is most likely to occur. Explain why?

Solution

Let

M be the event that the nurse is Male

F be the event that the nurse is Female

B be the event that the nurse has bachelor’ degree

B\` the event that the nurse does not have bachelor’s degree

- The probability distribution of recent study on nurse

Bachelor's Degree	Gender		Total
	Male (M)	Female (F)	
Has (B)	n(BM) = 34	n(BF) = 56	n(B) = 90
Doesn't (B`)	n(B`M) = 41	n(B`F) = 69	n(B`) = 110
Total	n(M) = 75	n(F) = 125	n(S) = 200

The probability that the nurse is

(a) P(FB) = n(FB) / n(S) = 56 / 200

(b) P(M) = n(M)/ n(S) = 75/ 200

(c) P(MB) = n(MB)/ n(S) = 34/ 200

(d) Part (b) is most likely to occur because its probability is more than a and c.

(e) The probability that the nurse is Male or he has bachelor's degree. P(M or B) = P(M) + P(B) –P(MB)

(f) P(M or F) = P(M) + P(F)

- Three cable channels (6, 8, and 10) have quiz shows, comedies, and dramas. The number of each is

Type of Show	TYPE OF CHANNEL			Total
	Channel 6	Channel 8	Channel 10	
Quiz show	5	2	1	8
Comedy	3	2	8	13
Drama	4	4	2	10
Total	12	8	11	31

If a show is selected at random, find the probabilities.

(a) The show is a quiz show, or it is shown on channel 8.

$$P(\text{ Q or C8}) = P(Q) + P(C8) - P(\text{Q and C8})$$

$$= 8/31 + 8/31 - 2/31 = 14/31$$

(b) The show is a Drama or a Comedy

$$P(\text{ D or C}) = P(D) + P(\text{ C}) = 10/31 + 13/31 = 23/\ 31$$

(c) The probability that the show is a shown on channel 10 or it is a drama.

P(C10 or D) = P(C10) + P(D) – P (D and C10)

= 11/31 + 10/31 – 2/31

= 19/31

(d) Are the type of channel and type of show are independent.

If the type of channel and type of show are independent

P(A and B) = P(A) *P(B)

P(Q and C6) = P(Q)* P(C6)=

L.H.S = P(Q and C6) = 5/31 = 0.1613

R. H.S = P(Q) * P(C6)= 8/31 * 12/31 = 0.09989

L.H .S # R.H S

Therefore the type of channel 6 and Quiz show are not independent.

P(Q and C8) = P(Q) * P(C8)

L. H S = P(Q and C8) = 2/31

- If 37 % of high school students said that they exercise regularly, find the probability that 5 randomly selected high school students will say that they exercise regularly. Would you answer this event likely or unlikely to occur? Explain your answer.

Solution

let

R be the event that the high school student they are exercise regularly

P(R) = 0.37

The probability that 5 randomly selected high school students will say that they exercise regularly = 0.37*.37*.37*.37*.37

= (0.37)^5

- Given a sample space with event A and B such that P(A)= 0.342, P(B) = 0.279, and P(A or B) = 0.601. Are A and B mutually exclusive ? Are A and B independent . Find P(A\B), P(not B) and P(A and B).

- P(A or B) = P(A) + P(B) (If Mutually exclusive)
- 0.601 = 0.342 + 0.279
- 0.601 # 0.621

Therefore event A and B are not mutually exclusive.

OR

- Solution

$P(A) = 0.342$, $P(B) = 0.279$, $P(A \text{ or } B) = 0.601$

If event A and B are not mutually exclusive

$P(A \text{ or } B) = P(A) + P(B) – P(A \text{ and } B)$

$0.601 \quad = 0.342 + 0.279 – P(A \text{ and } B)$

$0.601 \quad = 0.621 – P(A \text{ and } B)$

$P(A \text{ and } B) = 0.02$

Therefore, event A and B are not mutually exclusive. Because $P(A \text{ and } B)$ is not equal to zero.

If A and B are independent ,

P(A and B) = P(A) * P(B)

L. H. S = P(A and B) = 0.02

R. H. S = P(A) *P(B) = 0.342*0.279=0.0954

L.H S and R.H.S are not equal. Therefore, event A and B are not independent.

Find

P(A|B) = P(A and B)/ P(B) = 0.02/ 0.279 = 0.0717

P(not B) = 1- P(B) = 1- 0.279 = 0.721

P(A and B) = 0.02

- In a pizza restaurant , 95% of the customers order pizza. If 65% of the customers order pizza and a salad, find the probability that a customers who orders pizza will also order a salad.

$$P(S|O) = P(S \text{ and } O) / P(O) = 0.65/0.95$$

Yangon University of Economics
Department of Statistics
PGDRS.113
Probability Theory

Chapter 4
Probability and Counting Rules
Exercises

Selecting Colored Balls: Urn 1 contains 5 red balls and 3 black balls. Urn 2 contains 3 red balls and 1 black ball. Urn 3 contains 4 red balls and 2 black balls. If an urn is selected at random and a ball is drawn, find the probability it will be red.

Solution:

let

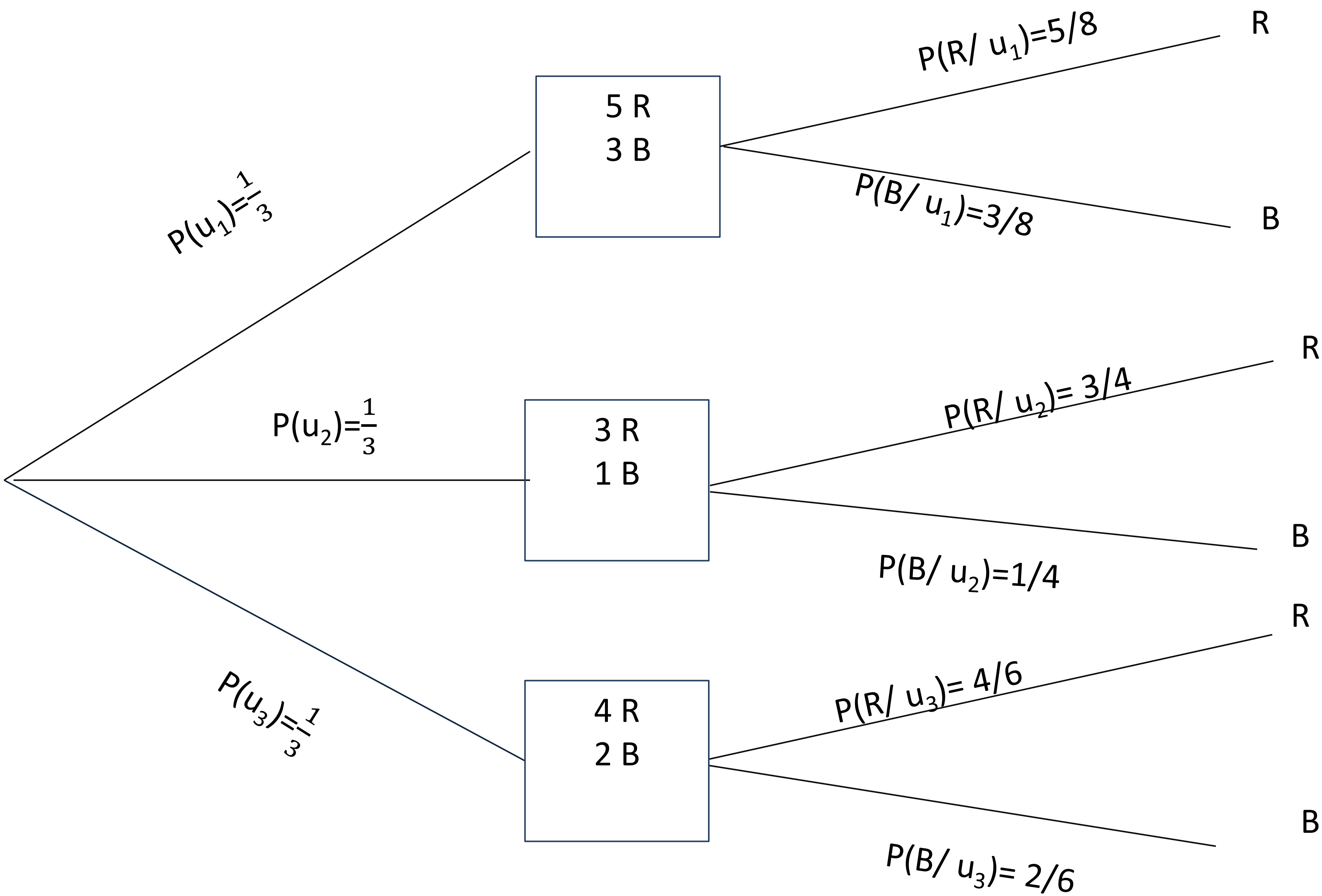
U_1 be the event that the selected urn is 1.

U_2 be the event that the selected urn is 2.

U_3 be the event that the selected urn is 3.

R be the event that a ball drawn is red.

B be the event that a ball drawn is black.



P(u1)=1/3
5 R
3 B
P(R/ u1)=5/8
R
P(B/ u1)=3/8
B
P(u2)=1/3
3 R
1 B
P(R/ u2)= 3/4
R
P(B/ u2)=1/4
B
P(u3)=1/3
4 R
2 B
P(R/ u3)= 4/6
R
P(B/ u3)= 2/6
B

The joint probabilities are

P(U_1 and R) = P(U_1)*P(R| U_1) = 1/3*5/8 = 5/24

P(U_1 and B) = P(U_1)*P(B| U_1) = 1/3*3/8 = 3/24

P(U_2 and R) = P(U_2)*P(R| U_2) = 1/3*3/4= 3/12

P(U_2 and B) = P(U_2)*P(B| U_2)= 1/3*1/4= 1/12

P(U_3 and R) = P(U_3)*P(R|U_3) = 1/3*4/6= 4/18

P(U_3 and B) = P(U_3)*P(B| U_3) = 1/3*2/6 = 2/18

The probability that a ball is drawn it will be red.

P(R)= P(U_1 and R)+ P(U_2 and R) + P(U_3 and R)

= 5/24 + 3/12 + 4/18

(i) Find the probability that it will be drawn red ball given that it is contain urn 1.

$P(R|U_1)$ = P(U1 and R) / P(U1) =

(ii) If the ball will be drawn red , the probability that the selected urn is 1.

$P(U_1 |R)$ = P(U_1 and R) / P(R) =

An insurance company classifies drivers low-risk, medium-risk, and high-risk. Of those insured, 60% are low-risk, 30% are medium-risk, and 10% are high-risk. After a study, the company finds that during a 1-year period, 1% of the low-risk drivers had accident, 5% of the medium-risk drivers had an accident, and 9% of the high-risk drivers had an accident. If a driver is selected at random, find the probability that the driver will have had an accident during the year.

Solution:

Let

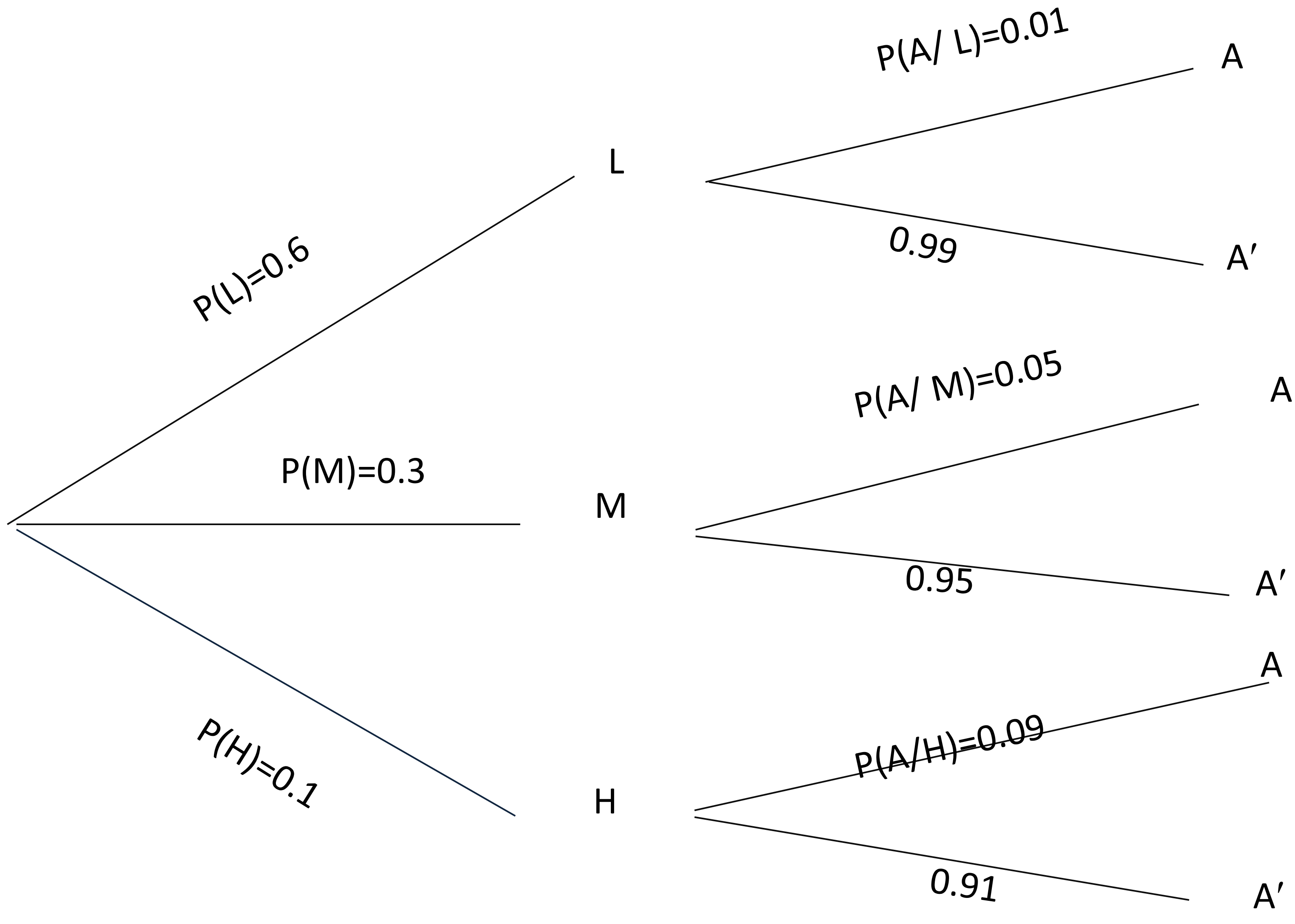
L be the event that the driver had low-risk insurance.

M be the event that the driver had medium-risk insurance.

H be the event that the driver had high-risk insurance.

A be the event that the driver had an accident during 1-year period.

A` be the event that the driver did not have an accident during 1- year period.


P(L)=0.6
P(M)=0.3
P(H)=0.1
L
M
H
P(A/ L)=0.01
0.99
P(A/ M)=0.05
0.95
P(A/H)=0.09
0.91
A
A′
A
A′
A
A′

The joint probability

$P(L \text{ and } A) = P(L)*P(A|L) = 0.6*0.01= 0.006$

$P(L \text{ and } A`) = P(L)*P(A`|L)= 0.6*0.99$

$P(M \text{ and } A) = P(M) * P(A|M) = 0.30 *0.05=0.015$

$P(M \text{ and } A`) = P(M) * P(A`|M) = 0.03 * 0.95$

$P(H \text{ and } A) = P(H) * P(A|H) = 0.1 * 0.09 = 0.009$

$P(H \text{ and } A`) = P(H) * P(A`|H) = 0.1 * 0.91$

The probability that the driver will have had an accident during the year.

$$
\begin{aligned}
P(A) &= P(L \text{ and } A) + P(M \text{ and } A) + P(H \text{ and } A) \\
&= 0.006 \quad + 0.015 \quad + 0.009 \\
&= 0.03(3\%)
\end{aligned}
$$

If the driver will have had an accident during the year, find the probability that he had high-risk insurance.

$P(H|A) = P(H \text{ and } A) / P(A) = 0.009 / 0.03 = 0.3$

Yangon University Of Economics

Department Of Statistics

PGDRS

Probability Theory
PGDRS 113

Presented By

Daw Aye Aye Htwe

Associated Professor

Department Of Statistics

"At Least" Probabilities

- Remember, complementary events should be used when finding "at least" probabilities
- Multiplication rule can be also be used for these types of events

Example 4-35

A person selects 3 cards from an ordinary deck and replaces each card after it is drawn. Find the probability that the person will get at least one heart.

Solution

Let E = *at least 1 heart is drawn and*

$\overline{E}$ = *no hearts are drawn*

$$P(E) = 1 - P(\overline{E})$$

$$= 1 - \left[\frac{39}{52} \cdot \frac{39}{52} \cdot \frac{39}{52}\right]$$

$$= \quad 1 - \left(\frac{39}{52}\right)^3$$

$$\approx 0.578 = 57.8\%$$

Hence, a person will select at least one heart about 57.8% of the time.

- “At least 1”
- P(At least 1) = 1- P(0)
- P(At least 2) = 1-[P(0)+P(1)]
- P(At least x) = 1- ($\sum_{i=0}^{n-1} P(n)$)

- P(A) + P(A’) = 1

  P(A) = 1- P(A’)

  P(Workers) = 0.9 P (no workers) = 0.1, n=3

  P(at least workers) = 1- P(no workers)

  = 1- (0.1*0.1*0.1)

  = 1- (0.1)^3

  P(at least Female) = 1- P(all males)

Example

- 4-36
  - A single die is rolled 4 times. Find the probability of getting at least one 6.

Solution

The probability of getting at least one 6.

$$P(\text{at least one } 6) = 1 - P(\text{no } 6s) = 1 - \left(\frac{5}{6}\right)^4$$

$$\approx 0.518$$

There is about a 51.8% chance of getting at least one 6 when a die is rolled four times.

Example

- 4-37

  The Neck-ware Association of America reported that 3% of ties sold in the USA are bow ties. If 4 customers who purchased ties are randomly selected, find the probability that at least 1 purchased a bow tie.

Solution

Let E = *at least 1 bow tie is purchased*

$\overline{E}$ = *no bow ties are purchased*

$P(E) = 0.03$

$P(\text{no bow ties are purchased}) = (0.97)(0.97)(0.97)(0.97) \approx 0.885;$

$P(\text{at least one bow tie is purchased}) = 1 - 0.885 = 0.115.$

There is an 11.5% chance of a person purchasing at least one bow tie.

4.4 – Counting Rules

- Many times, a person must know the number of all possible outcomes for a sequence of events

- Three rules can be used:

1. Fundamental counting rule
2. Permutation rule
3. Combination rule

Fundamental Counting Rule

- In a sequence of *n* events in which the first one has k_1 possibilities and the second event has k_2 possibilities, and so forth, the total number of possibilities of the sequence will be

$$\boldsymbol{k_1 * k_2 * k_3 * \cdots k_n}$$

- Note: In this case *and* means to multiply

Example

- 4-39
  - A paint manufacturer wishes to manufacture several different paints. The categories include

Color	Red, blue, white, black, green, brown, yellow
Type	Latex, oil
Texture	Flat, semigloss, high gloss
Use	Outdoor, indoor

Solution

You can choose one color and one type and one texture and one use. Since there are 7 color choices, 2 type choices, 3 texture choices, and 2 use choices, the total number of possible different paints is as follows:

Color	Type	Texture	Use	
7	2	3	2	= 84

When determining the number of different possibilities of a sequence of events, you must know whether repetitions are permissible.

Example

- 4–41

  The first year the state of Pennsylvania issued railroad memorial license plates, the plates had a picture of a steam engine followed by four digits. Assuming that repetitions

  are allowed, how many railroad memorial plates could be issued?

Solution

Since there are four spaces to fill for each space, the total number of plates that can be issued is

$$=10 \cdot 10 \cdot 10 \cdot 10 = 10{,}000.$$

Note: Actually there was such a demand for the plates, Pennsylvania had to use letters also.

Factorial Notation

- If repetitions are not permitted, then factorial notation can be used

- Factorial formula
  - For any counting n

$$n! = n * (n-1) * (n-2) * \cdots * 1$$

$$0! = 1$$

Permutation Rule

- **Permutation**
  - An arrangement of n objects in a specific order

- **Permutation Rule**
  - The arrangement of n objects in a specific order using r objects at a time is called a permutation of n objects taking r objects at a time. It is written as nPr and the formula is

$$nPr = \frac{n!}{(n-r)!}$$

Example

- 4–44

A radio talk show host can select 3 of 6 special guests for her program. The order of appearance of the guests is important. How many different ways can this be done?

Solution

Since the order of appearance on the show is important, there are ${}_6P_3$ *ways to select the* guests.

$${}_6P_3 = \frac{6!}{(6-3)!} = \frac{6!}{3!} = \frac{6.5.4.3.2.1}{3.2.1} = \frac{6.5.4.3!}{3!}$$

$$= 120$$

Hence, there can be 120 different ways to select 3 guests and present them on the Program in a specific order.

Example

- 4–45

A school musical director can select 2 musical plays to present next year. One will be presented in the fall, and one will be presented in the spring. If she has 9 to pick from, how many different possibilities are there?

Solution

Order is important since one play can be presented in the fall and the other play in the spring.

$$ {}_9P_2 = \frac{9!}{(9-2)!} = \frac{9!}{7!} = \frac{9.8.7!}{7!} $$

$= 72$

There are 72 different possibilities.

Example

- 4–46

  How many permutations of the letters can be made from the word *STATISTICS?*

Solution

In the word *STATISTICS, there are 3 S's, 3 T's, 2 I's, 1 A, and 1 C.*

$$\frac{10!}{3!3!2!1!1!} = \frac{10.9.8.7.6.5.4.3.2.1}{3.2.1.3.2.1.2.1.1.1}$$

$= 50{,}400$

There are 50,400 permutations that can be made from the word *STATISTICS.*

Combination Rule

- **Combination**
  - Selection of distinct objects **without regard to order**

- **Combination Rule**
  - The number of combinations of r objects selected from n objects is denoted by nCr and is given by the formula

$$\boldsymbol{nCr = \frac{n!}{r!(n-r)!}}$$

Combination

A selection of distinct objects without regard to order is called a combination.

The difference between a combination and a permutation can be shown using the letters A, B, C, and D. what is left is a list of combinations, as shown.

AB	BC CD
AC	BD
AD	

Hence, the combinations of A, B, C, and D are AB, AC, AD, BC, BD, and CD. (Alternatively, BA could be listed and AB crossed out, etc.) The combinations have been listed alphabetically for convenience, but this is not a requirement.

The difference between a combination and a permutation can be shown using the letters A, B, C, and D. The permutations for the letters A, B, C, and D are

AB	BA	CA	DA
AC	BC	CB	DB
AD	BD	CD	DC

In permutations, AB is different from BA. But in combinations, AB is the same as BA since the order of the objects does not matter in combinations. Therefore, if duplicates are removed from a list of permutations,

Example

- 4-47
  - How many combinations of 4 objects are there, taken 2 at a time?

Solution

$$ {}_4C_2 = \frac{4!}{(4-2)!2!} = \frac{4.3.2!}{2.1.2!} = 6 $$

Example

- 4–48

  The director of Movies at the Park must select 4 movies from a total of 10 movies to show on Movie Night at the Park. How many different ways can the selections be made?

Solution

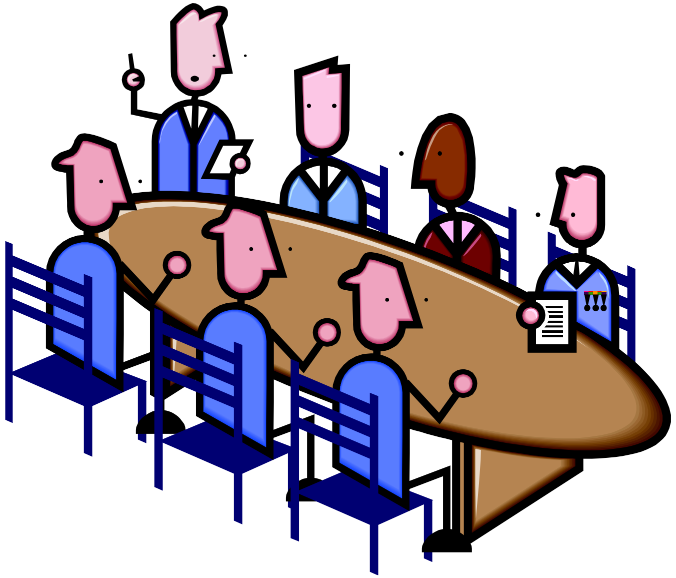
$$ {}_{10}C_4 = \frac{10!}{(10-4)!4!} = \frac{10!}{6!4!} $$

= *210*

The director has 210 different ways to select four movies from 10 movies. In this case, the order in which the movies are shown is not important.

Examples

- 4-49
  - In a club there are 7 women and 5 men. A committee of 3 women and 2 men is to be chosen. How many different possibilities are there?



Solution

We select 3 women from 7 women, which can be done in ${}_{7}C_{3}$, *or 35*, *ways*. Next, 2 men must be selected from 5 men, which can be done in ${}_{5}C_{2}$, *or 10*, *ways*. Finally, by the fundamental counting rule, the total number of different ways is 35 · 10 = 350, since you are choosing both men and women.

$$7C3 \cdot 5C2 = \frac{7!}{(7-3)!3!} \cdot \frac{5!}{(5-2)!2!} = 350$$

4.5 – Probability & Counting Rules

- Counting rules can be combined with probability rules to solve many types of probability problems

- Example 4-50
  - Find the probability of getting 4 Aces when 5 cards are drawn from an ordinary deck of cards.

Solution

There are ${}_{52}C_5$ *ways to draw 5 cards from a deck. There is only 1 way to get 4 aces (that* is, ${}_{4}C_{4}$*), but there are 48 possibilities to get the fifth card. Therefore, there are 48 ways* to get 4 aces and 1 other card. Hence,

$$P(4\ aces) = \frac{{}_{4}C_{4}.48}{{}_{52}C_{5}} = \frac{48}{2598960} = \frac{1}{54145}$$

Example

- 4-51
  - A box contains 24 integrated circuits, 4 of which are defective. If 4 are sold at random, find the following probabilities

a. Exactly 2 are defective

b. None is defective

c. All are defective

d. At least 1 is defective

Solution

a) *P(exactly 2 defectives)=* $\dfrac{{}_{4}C_{2}\cdot{}_{20}C_{2}}{{}_{24}C_{4}}=\dfrac{190}{1771}$

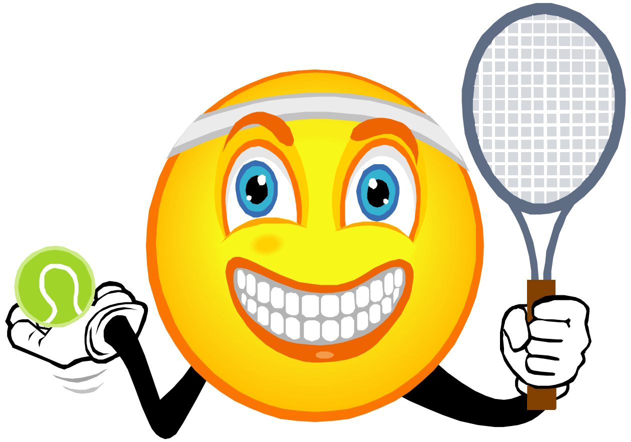
b) *P(no defectives) =* $\dfrac{{}_{20}C_{4}}{{}_{24}C_{4}}=\dfrac{1615}{3542}$

c) *P(all defective) =* $\dfrac{1}{{}_{24}C_{4}}=\dfrac{1}{10626}$

d) *P(at least 1 defective) =* $1-\dfrac{{}_{20}C_{4}}{{}_{24}C_{4}}=\dfrac{1927}{3542}$

Example

- 4-54
  - There are 8 married couples in a tennis club. If 1 man and 1 woman are selected at random to plan the summer tournament, find the probability that they are married to each other.



Solution

Since there are 8 ways to select the man and 8 ways to select the woman, there are 8 · 8, or 64, ways to select 1 man and 1 woman. Since there are 8 married couples, the solution

is $\frac{8}{64} = \frac{1}{8}$.

Yangon University Of Economics

PGDRS

Department Of Statistics

Probability Theory
PGDRS 113

Presented By

Daw Aye Aye Htwe

Associated Professor

Department Of Statistics

EXAMPLE 5–8 On Hold for Talk Radio

A talk radio station has four telephone lines. If the host is unable to talk (i.e., during a commercial) or is talking to a person, the other callers are placed on hold. When all lines are in use, others who are trying to call in get a busy signal.

The probability that variance and standard deviation for the distribution.

	0	1	2	3	4
P(X)	0.18	0.34	0.23	0.21	0.04

Should the station have considered getting more phone lines installed?

S O LU T I O N

The mean is

$$\mu = \sum X.P(X)$$

=0*(0.18)+1*(0.34)+2*(0.23)+3*(0.21)+4*(0.04) =1.59

The variance is

$$\sigma^2 = \sum \left[X^2.P(X)\right] - \mu^2$$

$$= [(\mathbf{0})^{2} * \mathbf{0.18} + (\mathbf{1})^{2} * \mathbf{0.34} + (\mathbf{2})^{2} * \mathbf{0.23} + (\mathbf{3})^{2} * \mathbf{0.21} + (\mathbf{4})^{2} * \mathbf{0.04}] = \mathbf{1.262}]$$

The standard deviation $\sigma = \sqrt{\sigma^2}$ (or)

$$\sigma = \sqrt{1.262} \approx 1.121$$

No. The mean number of people calling at any one time is 1.59. Since the standard deviation is 1.123, most callers would be accommodated by having four phone lines

Because $\mu + 2\sigma$ would be $1.59 + 2(1.123) = 3.836 \approx 4.0$. Very few callers would get a busy signal since at least 75% of the callers would either get through or be put on hold.

Expectation

Another concept related to the mean for a probability distribution is that of expected value or expectation. Expected value is used in various types of games of chance, in insurance, and in other areas, such as decision theory.

The **expected value** of a discrete random variable of a probability distribution is the

theoretical average of the variable. The formula is

$$\mu = E(X) = \Sigma X \cdot P(X)$$

The symbol E(X) is used for the expected value. The formula for the expected value is the same as the formula for the theoretical mean. The expected value, then, is the theoretical mean of the probability distribution.

That is, $E(X) = \mu$. When expected value problems involve money, it is customary to round the answer to the nearest cent.

EXAMPLE 5-9 Winning Tickets

One thousand tickets are sold at $1 each for a smart television valued at $750. What is the expected value of the gain if you purchase one ticket?

SOLUTION

The problem can be set up as follows:

Gain X	$749	−$1
Probability P(X)	1/1000	999/1000

Two things should be noted. First, for a win, the net gain is $749, since you do not get the cost of the ticket ($1) back. Second, for a loss, the gain is represented by a negative number, in this case −$1. The solution, then, is

X	P(X)	X P(X)
Gain $749	$\frac{1}{1000}$	$\frac{749}{1000}$
Loss (-$1)	$\frac{999}{1000}$	$-\frac{999}{1000}$
		$-\frac{25}{1000} = -\$\,0.25$

$E(X) = (\$749)*1/1000 + (-\$1)*999/1000 = -\$0.25$

Hence, a person would lose, on average, −\$0.25 on each ticket purchased.

Expected value problems of this type can also be solved by finding the overall gain(i.e., the value of the prize won or the amount of money won, not considering the cost of the ticket for the prize or the cost to play the game) and subtracting the cost of the tickets or the cost to play the game, as shown:

$E(X) = (\$750)*1/1000 - \$1 = -\$0.25$

Note that the expectation is −\$0.25. This does not mean that you lose \$0.25, since you can only win a television set valued at \$750 or lose \$1 on the ticket. What this expectation means is that the average of the losses is \$0.25 for each of the 1000 ticket holders. Here is another way of looking at this situation: If you purchased one ticket each week over a long time, the average loss would be \$0.25 per ticket, since theoretically, on average, you would win the television set once for each 1000 tickets purchased.

EXAMPLE 5–10 Bond Investment

A financial adviser suggests that his client select one of two types of bonds in which to invest \$5000. Bond *X* pays a return of 4% and has a default rate of 2%. Bond *Y* has a 2.5 % return and a default rate of 1%. Find the expected rate of return and decide which bond would be a better investment. When the bond defaults, the investor loses all the investment.

SOLUTION

The return on bond *X* is \$5000 · 4% = \$200. The expected return then is

$$E(X) = \$200(0.98) - \$5000(0.02) = \$96$$

The return on bond *Y* is \$5000 · 2 .5% = \$125. The expected return then is

$$E(Y) = \$125(0.99) - \$5000(0.01) = \$73.75$$

Hence, bond *X* would be a better investment since the expected return is higher.

Exercise

1. **Suit Sales :**The number of suits sold per day at a retail store is shown in the table, with the corresponding probabilities. Find the mean, variance, and standard deviation of the distribution.

Number of suits sold *X*	19	20	21	22	23
Probability *P*(*X*)	0.2	0.2	0.3	0.2	0.1

If the manager of the retail store wants to be sure that he has enough suits for the next 5 days, how many should the manager purchase?

X	P(X)	X P(X)	X^2 P(X)
19	0.2	3.8	72.2
20	0.2	4	80
21	0.3	6.3	132.3
22	0.2	4.4	96.8
23	0.1	2.3	52.9
		20.8	434.2

- Mean, E(X)=20.8 = 21suits
- Variance , V(X)= 434.2- (20.8)^2 = 1.56 (Suits)^2
- Standard Deviation , SD(X) = $\sqrt{1.56}$ = 1.249 (suits)
- For the next 5 days , he would purchases,

$$= 5 * 20.8 = 104 \text{ suits}$$

2. Winning the Lottery:

For a daily lottery, a person selects a three- digit number. If the person plays for $1, she can win $500. Find the expectation.

$E(X) = \$\, 499 * 1/1000 + (-\$1) * 999/1000 = -\ \$\, 0.5$

In the same daily lottery. If a person boxes a number , she will win $80. Find the expectation if the number 123 is played for $1 and boxed.

$E(X) = \$\, 79 * 1/1000 + (-\ \$\, 1) * 999/1000 = -\$\, 0.92$

Yangon University Of Economics
Department of Statistics
PGDRS

PGDRS-113

Probability Theory

Week 7

Probability Theory
PGDRS 113

Presented By

Daw Aye Aye Htwe

Associated Professor

Department Of Statistics

5–4 Other Types of Distributions

In addition to the binomial distribution, other types of distributions are used in statistics. Four of the most commonly used distributions are the multinomial distribution, the Poisson.

The Multinomial Distribution

Recall that for an experiment to be binomial, two outcomes are required for each trial. But if each trial in an experiment has more than two outcomes, a distribution called the **multinomial distribution** must be used. For example, a survey might require the responses of “approve,” “disapprove,” or “no opinion.” In another situation, a person may have a choice of one of five activities for Friday night, such as a movie, dinner, baseball game, play, or party. Since these situations have more than two possible outcomes for each trial, the binomial distribution cannot be used to compute probabilities.

A **multinomial experiment** is a probability experiment that satisfies the following four requirements:

1. There must be a fixed number of trials.

2. Each trial has a specific—but not necessarily the same—number of outcomes.

3. The trials are independent.

4. The probability of a particular outcome remains the same.

Formula for the Multinomial Distribution

$$P(X) = \frac{n!}{X_1!.X_2!.X_3!\ldots X_k!}.P_1^{x_1}.P_2^{x_2}\ldots P_k^{x_k}$$

Where $X_1 + X_2 + X_3 + \ldots + X_k = n$ and $P_1 + P_2 + P_3 + \ldots + P_k = 1$

EXAMPLE 5–24 Coffee Shop Customers

A small airport coffee shop manager found that the probabilities a customer buys 0,1,2, or 3 cups of coffee are 0.3.0.5,0.15, and 0.05, respectively. If 8 customers enter the shop, find the probability that 2 will purchase something other than coffee ,4 will purchase 1 cup of coffee, 1 will purchase 2cups, and 1 will purchase 3 cups.

SOLUTION

Let

$$n = 8, \qquad X_1 = 2, X_2 = 4, X_3 = 1, X_4 = 1; \qquad P_1 = 0.3, P_2 = 0.5, P_3 = 0.15, and P_4 = 0.05$$

$$P(X) = \frac{8!}{2_1!.4_2!.1_3!1!}(0.3)^2(0.5)^4(0.15)^1(0.05)^1 \approx 0.0354$$

There is a 0.0354 probability that the results will occur as described.

EXAMPLE 5–25 Herbicides

It was found that 65% of individuals use herbicides for commercial purposes, 27% of individuals use herbicides for agricultural purposes, and 8% of individuals use herbicides for home and garden purposes. Of 5 people who said that they used herbicides, find the probability that 3 used them for commercial purposes, 1 used them for agriculture purposes, and 1 used them for home or garden purposes.

Source: EPA.

SOLUTION

Let $n = 5$, $X_1 = 3$, $X_2 = 1$, $X_3 = 1$, $p_1 = 0.65$, $p_2 = 0.27$, and $p_3 = 0.08$. Substituting in the formula gives

$$P(X) = \frac{5!}{3! \cdot 1! \cdot 1!} \cdot (0.65)^3(0.27)^1(0.08)^1 \approx 0.119$$

There is a 0.119 probability that if 5 people are selected, 3 will use herbicides for commercial purposes, 1 person will use them for agricultural purposes, and 1 person will use them for home and garden purposes.

The Poisson Distribution

A discrete probability distribution that is useful when n is large and p is small and when the independent variables occur over a period of time is called the **Poisson distribution.** In addition to being used for the stated conditions (that is, n is large, p is small, and the variables occur over a period of time), the Poisson distribution can be used when a density of items is distributed over a given area or volume, such as the number of plants growing per acre or the number of defects in a given length of videotape.

A **Poisson experiment** is a probability experiment that satisfies the following requirements:

1. The random variable X is the number of occurrences of an event over ., length, area, volume, period of time, etc.).

2. The occurrences occur randomly.

3. The occurrences are independent of one another.

4. The average number of occurrences over an interval is known.

Formula for the Poisson Distribution

The probability of X occurrences in an interval of time, volume , area ,etc ., for variable where is the mean number of occurrence per unit (λ) is

$$P(X;\lambda)=\frac{e^{-\lambda}\lambda^{x}}{X!};X=0,1,2,...$$

The letter *e* is a constant approximately equal to 2.7183.

EXAMPLE 5–25 Typographical Errors

If there are 200 typographical errors randomly distributed in a 500-page manuscript, find the probability that a given page contains exactly 3 errors.

S O LU T I O N

First, find the mean number λ of errors. Since there are 200 errors distributed over 500 pages, each page has an average of

λ =200/500= 2/5= 0.4

or 0.4 error per page. Since *X* = 3, substituting into the formula yields

$$P(X;\lambda) = \frac{e^{-\lambda}\ \lambda^{0.4}}{X!} = \frac{e^{-0.4}\ 0.4^3}{3!} \approx 0.0072$$

Thus, there is less than a 1% chance that any given page will contain exactly 3 errors.

EXAMPLE 5–26 Toll-Free Telephone Calls

A sale firm receives, on average, 3 calls per hour on its toll-free number. For any given hour, find the probability that it will receive the following.

(a) At most 3 calls (b) At least 3 calls(c) 5 or more calls

S O LU T I O N

Let X be the number of telephone call per hour.

λ be the average of telephone call per hour.

$$P(X;\lambda) = \frac{e^{-\lambda}\lambda^{x}}{X!} : X = 0,1,2,...$$

$$P(X;3) \quad = \frac{e^{-3}3^{x}}{x!} : x = 0,1,2,\dots\dots\dots$$

$$P(X=0;\lambda=3) = \frac{e^{-3}3^{0}}{0!} = 0.0498$$

SOLUTION

$$P(X;\lambda)=\frac{e^{-\lambda}\lambda^{x}}{X!}:X=0,1,2,\ldots$$

$$P(X=0;\lambda=3)=\frac{e^{-3}3^{0}}{0!}=0.0498$$

$$P(X=1;\lambda=3)=\frac{e^{-3}3^{1}}{1!}=0.1494$$

$$P(X=2;\lambda=3)=\frac{e^{-3}3^{2}}{2!}=0.2240$$

$$P(X=3;\lambda=3)=\frac{e^{-3}3^{3}}{3!}=0.2240$$

$$P(X=4;\lambda=3)=\frac{e^{-3}3^{4}}{4!}=0.1680$$

a. “At most 3 calls” means 0, 1, 2, or 3 calls. Hence,

$$P(x \leqslant 3) = P(x=0)+P(x=1)+P(x=2)+P(x=3)$$
$$= 0.0498+0.1494+0.2240+0.2240 = 0.6472$$

b. “At least 3 calls” means 3 or more calls. It is easier to find the probability of 0, 1, and 2 calls and then subtract this answer from 1 to get the probability of at least 3 calls.

$$P(x \geqslant 3) = P(x=3)+P(x=4)+P(x=5)+\ldots\ldots$$

(or)

$$= 1-P(x<3)$$
$$= 1-[P(x=0)+P(x=1)+P(x=2)]$$
$$= 1-(0.0498+0.1494+0.2240)$$
$$= 1-0.4232 = 0.5768$$

c. For the probability of 5 or more calls, it is easier to find the probability of getting 0, 1, 2, 3, or 4 calls and subtract this answer from 1. Hence,

$P(x \geqslant 5) = P(x=5)+P(x=6)+\ldots.$

(or)

$= 1-P(x<5)$

$= 1- [P(x=0)+P(x=1)+P(x=2)+P(x=3)+P(x=4)]$

$= 1-(0.0498+0.1494+0.2240+0.2240+0.1680)$

$= 1-0.8152=0.1848$

Thus, for the event described, the part a events is most likely to occur, and the part c event is least likely to occur.

Binomial Approximation

- The Poisson distribution can also be used to approximate the binomial distribution when the expected value $\lambda = n \cdot p$ is less than 5, as shown in Example 5–30. (The same is true when $n \cdot q < 5$.)

- In Binomial Distribution , Mean is $\mu = n*p$
- In Poisson Distribution, Mean is λ
- Mean , $\lambda = \mu = n*p$ (less than 5)

Mean, Variance and Standard Deviation

Mean

$\mu = \lambda$

Variance

$\sigma^2 = \lambda$

Standard Deviation

$\sigma = \sqrt{\lambda}$

EXAMPLE 5–30 Left-Handed People

If approximately 2% of the people in a room of 200 people are left-handed, find the probability that exactly 5 people there are left-handed.

SOLUTION

Since $\lambda = n \cdot p$, then $\lambda = (200)(0.02) = 4$. Hence,

$$P(X; \lambda) = \frac{(2.7183)^{-4}(4)^5}{5!} \approx 0.1563$$

which is verified by the formula ${}_{200}C_5(0.02)^5(0.98)^{195} \approx 0.1579$. The difference between the two answers is based on the fact that the Poisson distribution is an approximation and rounding has been used.


THANK YOU
& GOODNIGHT

Yangon University Of Economics
Department of Statistics
PGDRS

PGDRS-113

Probability Theory

Week 8

Probability Theory
PGDRS 113

Presented By

Daw Aye Aye Htwe

Associated Professor

Department Of Statistics

The Hyper-geometric Distribution

When sampling is done *without* replacement, the binomial distribution does not give exact probabilities, since the trials are not independent. The smaller the size of the population, the less accurate the binomial probabilities will be.

For example, suppose a committee of 4 people is to be selected from 7 women and 5 men. What is the probability that the committee will consist of 3 women and 1 man?

- To solve this problem, you must find the number of ways a committee of 3 women and 1 man can be selected from 7 women and 5 men. This answer can be found by using combinations; it is

$$ {}_7C_3 \cdot {}_5C_1 = 175 $$

- Next, find the total number of ways a committee of 4 people can be selected from 12 people. Again, by the use of combinations, the answer is

$$ {}_{12}C_4 = 495 $$

Finally, the probability of getting a committee of 3 women and 1 man from 7 women and 5 men is

$P(X) = 175/495 = 35/99 \; = 0.3535$

The results of the problem can be generalized by using a special probability distribution called the hyper-geometric distribution. The **hyper-geometric distribution** is a distribution of a variable that has two outcomes when sampling is done without replacement.

A **hyper-geometric experiment** is a probability experiment that satisfies the following requirements:

1. There are a fixed number of trials.

2. There are two outcomes, and they can be classified as success or failure.

3. The sample is selected without replacement

Formula for the Hyper-geometric Distribution

Given a population with only two types of objects (females and males, defective and non defective, successes and failures, etc.), such that there are *a* items of one kind and *b* items of another kind and $a + b$ equals the total population, the probability $P(X)$ of selecting without replacement a sample of size *n* with *X* items of type *a* and

$n - X$ items of type *b* is $P(X) = \dfrac{{}_aC_x \cdot {}_bC_{n-x}}{{}_{a+b}C_n}$

$$P(X) = \frac{{}_aC_x \cdot {}_bC_{n-x}}{{}_{a+b}C_n}$$

The basis of the formula is that there are ${}_aC_x$ ways of selecting the first type of items, ${}_bC_{n-x}$ ways of selecting the second type of items, and ${}_{a+b}C_n$ ways of selecting *n* items from the entire population.

EXAMPLE 5–27 Assistant Manager Applicants

Ten people apply for a job as assistant manager of a restaurant. Five have completed college and five have not. If the manager selects 3 applicants at random, find the probability that all 3 are college graduates.

SOLUTION

Assigning the values to the variables gives

a = 5 college graduates, $n = 3$, $X = 3$

b = 5 non-graduates ,

and $n - X = 0$. Substituting in the formula gives

$$P(X) = \frac{{}_aC_x \cdot {}_bC_{n-x}}{{}_{a+b}C_n} = \frac{{}_5C_3 \cdot {}_5C_0}{{}_{10}C_3} = \frac{10}{120} = \frac{1}{12} \approx 0.083$$

There is a 0.083 probability that all 3 applicants will be college graduates.

EXAMPLE 5–32 House Insurance

A recent study found that 2 out of every 10 houses in a neighborhood have no insurance. If 5 houses are selected from 10 houses, find the probability that exactly 1 will be uninsured.

SOLUTION

In this example, $a = 2$, $b = 8$, $n = 5$, $X = 1$, and $n - X = 4$.

$$P(X) = \frac{{}_2C_1 \cdot {}_8C_4}{{}_{10}C_5} = \frac{2 \cdot 70}{252} = \frac{140}{252} = \frac{5}{9} \approx 0.556$$

There is a 0.556 probability that out of 5 houses, 1 house will be uninsured.

Geometric Distribution

A geometric experiment is a probability experiment if it satisfies the following requirements:

- 1. Each trial has two outcomes that can be either success or failure.
- 2. The outcomes are independent of each other.
- 3. The probability of a success is the same for each trial.
- 4. The experiment continues until a success is obtained.

Formula for the Geometric Distribution

- If p is the probability of a success on each trial of a binomial experiment and n is the number of the trial at which the first success occurs, then the probability of getting the first success on the nth trial is

$$P(n) = p(1 - p)^{n-1}$$

- where $n = 1, 2, 3, \ldots$

EXAMPLE 5–35 Blood Types

In the United States, approximately 42% of people have type A blood. If 4 people are selected at random, find the probability that the fourth person is the first one selected with type A blood.

SOLUTION

Let $p = 0.42$ and $n = 4$.

$$P(n) = p(1 - p)^{n-1}$$

$$P(4) = (0.42)(1 - 0.42)^{4-1}$$

$$= (0.42)(0.58)^3$$

$$\approx 0.0819 \approx 0.082$$

There is a 0.082 probability that the fourth person selected will be the first one to have type A blood.


Good
Night


Sweet
Dreams


THANK YOU
FOR YOUR
ATTENTION

YANGON UNIVERSITY OF ECONOMICS

DEPARTMENT OF STATISTICS

PGDRS

Probability Theory
PGDRS 113

Presented By

Daw Aye Aye Htwe

Associated Professor

Department Of Statistics

CHECKING FOR NORMALITY

Histogram

Pearson’s Index PI of Skewness

Outliers

Other Tests

- Normal Quantile Plot
- Chi-Square Goodness-of-Fit Test
- Kolmogorov-Smikirov Test
- Lilliefors Test

DETERMINING NORMALITY

Histogram

To draw a histogram for the data and check its shape. If the histogram is not approximately bell- shaped, then the data are not normally distributed.

DETERMINING NORMALITY

Pearson’s Index PI of Skewness

Skewness can be checked by using the Pearson coefficient (PC) of skewness also called Pearson’s index of skewness. The formula is

$$PC = \frac{3(\bar{X} - median)}{s}$$

If the index is greater than or equal to +1 or less than or equal to −1, it can be concluded that the data are significantly skewed.

DETERMINING NORMALITY

Outliers

Q_1 , Q_2 , Q_3, $IQR = Q_3 - Q_1$

the data should be checked for outliers by using the method

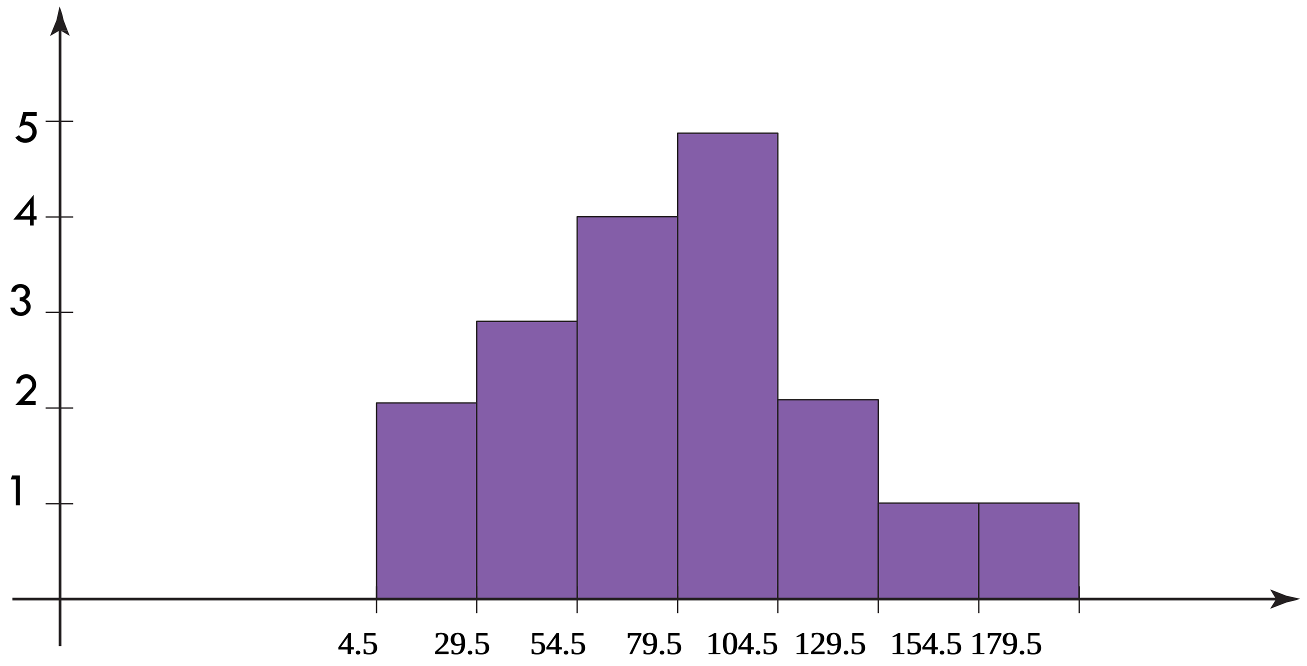
EXAMPLE 6–11 TECHNOLOGY INVENTORIES

A survey of 18 high-tech firms showed the number of days' inventory they had on hand. Determine if the data are approximately normally distributed.

5	29	34	44	45	63	68	74	74
81	88	91	97	98	113	118	151	158

Solution Frequency distribution

Class limits	Frequency	class boundaries
5–29	2	4.5 – 29.5
30–54	3	29.5-54.5
55–79	4	54.5- 79.5
80–104	5	79.5-104.5
105–129	2	104.5-129.5
130–154	1	129.5-154.5
155–179	1	154.5 – 179.5
Total	18	



Since the histogram is approximately bell-shaped, we can say that the distribution is approximately normal

$$\bar{X} = \frac{\sum x}{n} = \frac{1431}{18} = 79.5$$

$$\text{Median} = \frac{74+81}{2} = 77.5$$

$$S = \frac{\sum x^2 - n\bar{x}^2}{n-1} = \frac{\sum 141585 - 18 \times 79.5^2}{17} = 40.5$$

$$PC = \frac{3(\bar{X} - median)}{s}$$

$$= \frac{3(79.5 - 77.5)}{40.5}$$

$$= 0.148$$

PC is not greater than +1 or less than −1, so it can be concluded that the distribution is not significantly skewed.

$Q1 = 45$ and $Q3 = 98$

$IQR = Q3 - Q1 = 98 - 45 = 53$

$45 - 1.5(53) = -34.5$

$98 + 1.5(53) = 177.5$

There are no outliers.

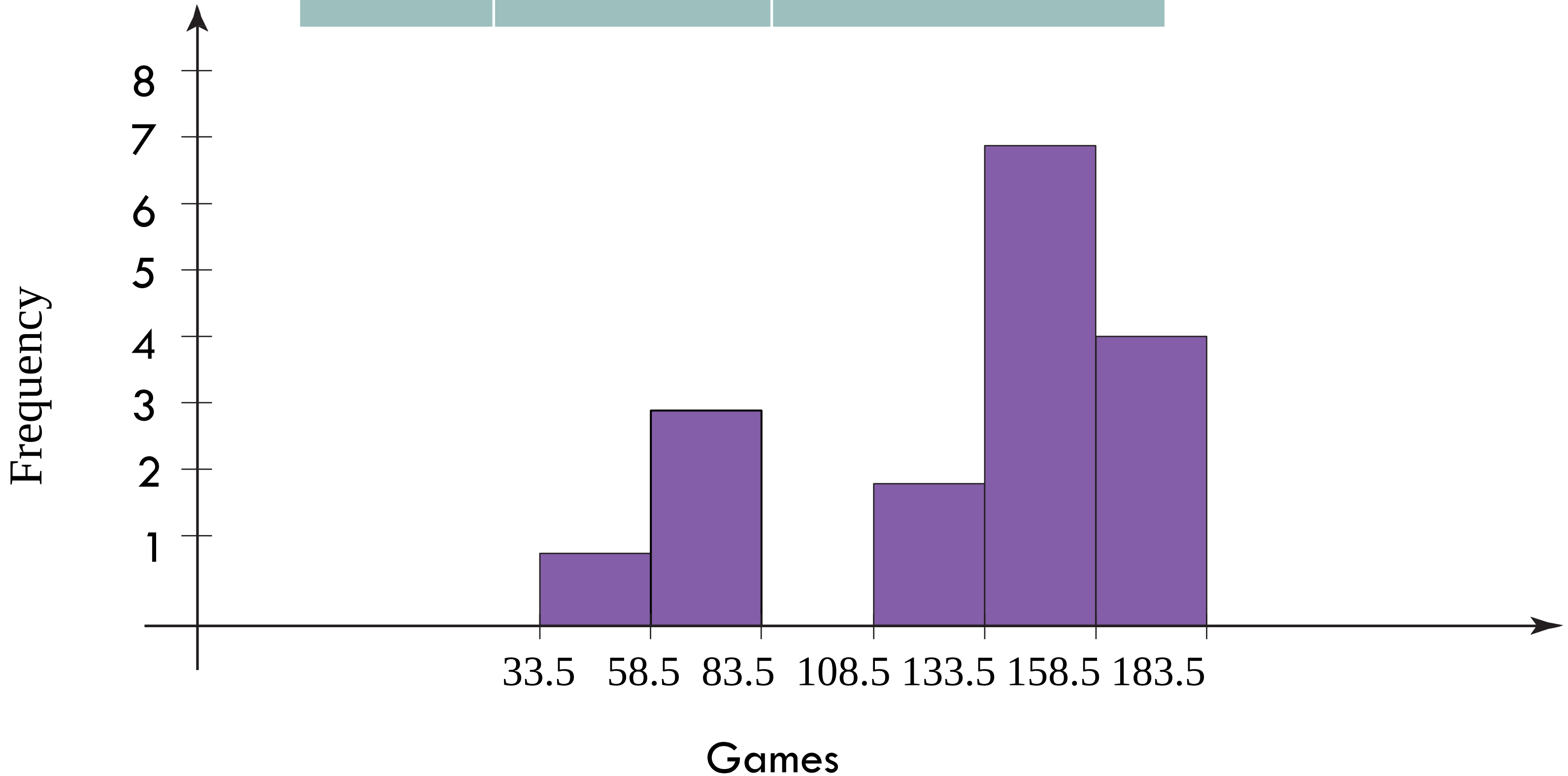
Since the histogram is approximately bell-shaped, the data are not significantly skewed, and there are no outliers, it can be concluded that the distribution is approximately normally distributed.

EXAMPLE 6–12NUMBER OF BASEBALL GAMES PLAYED

The data shown consist of the number of games played each year in the career of Base-ball Hall of Famer Bill Mazeroski. Determine if the data are approximately normally distributed.

81 148 152 135 151 152 159 142 34 162 130 162
163 143 67 112 70

Class	Frequency	Class boundaries
34–58	1	33.5-58.5
59–83	3	58.5-83.5
84–108	0	83.5-108.5
109–133	2	108.5-133.5
134–158	7	133.5-158.5
159–183	4	158.5-183



X = 127.24, median = 143, and s = 39.87.

$$PC = \frac{3(\bar{X} - median)}{s}$$

$= 3(127.24 - 143)$

$= 39.87$

≈ -1.19

Since the PC is less than −1, it can be concluded that the distribution is significantly skewed to the left.

Q1 = 96.5 and Q3 = 155.5.

IQR = Q3 − Q1 = 155.5 − 96.5 = 59

96.5 − 1.5(59) = 8

155.5 + 1.5(59) = 244

There are no outliers. The distribution is negatively skewed.

6.3 THE CENTRAL LIMIT THEOREM

In addition to knowing how individual data values vary about the **mean** for a **population**, statisticians are interested in knowing how the means of samples of the same size taken from the same population vary about the population mean.

DISTRIBUTION OF SAMPLE MEANS

A ***sampling distribution of sample means*** is a distribution obtained by using the means computed from random samples of a specific size taken from a population.

Sampling error is the difference between the sample measure and the corresponding population measure due to the fact that the sample is not a perfect representation of the population.

PROPERTIES OF THE DISTRIBUTION OF SAMPLE MEANS

- The mean of the sample means will be the same as the population mean.

- The standard deviation of the sample means will be smaller than the standard deviation of the population, and will be equal to the population standard deviation divided by the square root of the sample size.

WITH REPLACEMENT

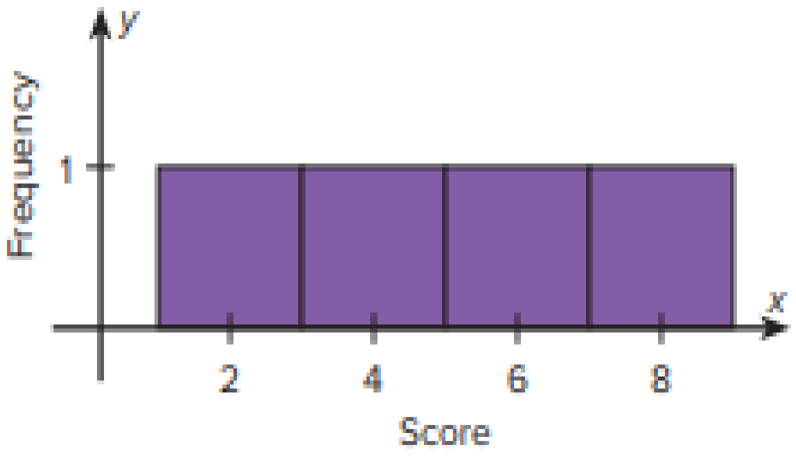
The following example illustrates these two properties.

Suppose a professor gave an 8-point quiz to a small class of four students. The results of the quiz were 2, 6, 4, and 8. For the sake of discussion, assume that the four students constitute the population.

X: 2,6,4,8   N=4

The mean of the population is

$$\mu = \frac{2+6+4+8}{4} = 5$$



The standard deviation of the population is

$$\sigma = \frac{\sqrt{(2-5)^2 + (6-5)^2 + (4-5)^2 + (8-5)^2}}{4} \approx 2.236$$

The graph of the original distribution is shown in Figure 6–31. This is called a *uniform distribution.*

WITH REPLACEMENT

The number of possible sample = $N^n = 4^2 = 16$

1. $\mu_{\overline{x}} = \mu$

The mean of the sampling distribution is unbiased estimator of the population.

2. $\sigma_{\overline{x}}^2 = \dfrac{\sigma^2}{n}$ OR $\sigma_{\overline{x}} = \dfrac{\sigma}{\sqrt{n}}$

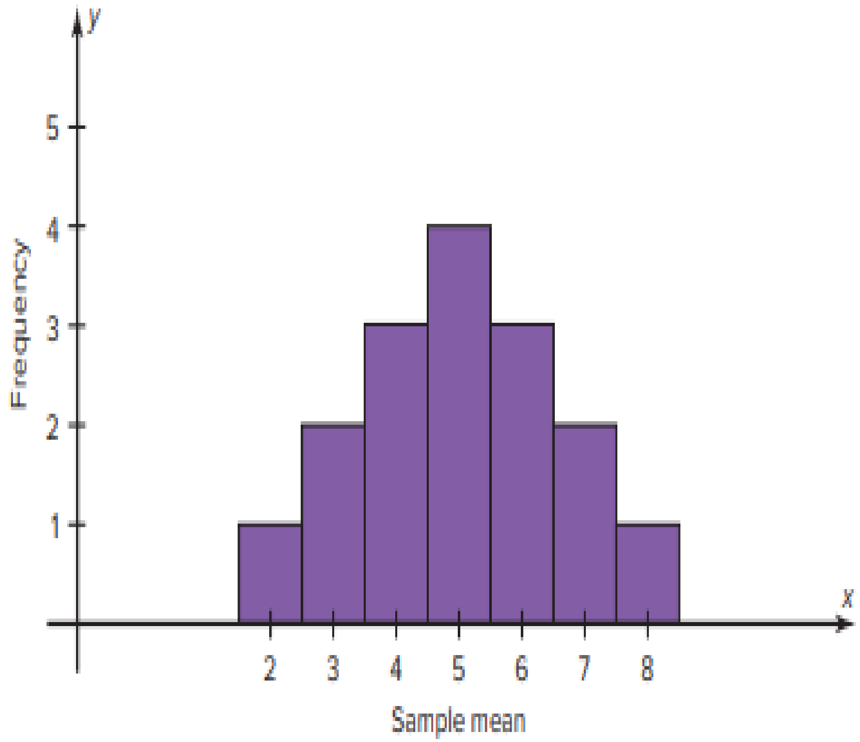
Now, if all samples of size 2 are taken with replacement and the mean of each sample is found, the distribution is as shown.

Sample	Mean	Sample	Mean
2, 2	2	6, 2	4
2, 4	3	6, 4	5
2, 6	4	6, 6	6
2, 8	5	6, 8	7
4, 2	3	8, 2	5
4, 4	4	8, 4	6
4, 6	5	8, 6	7
4, 8	6	8, 8	8

SAMPLING DISTRIBUTION OF THE SAMPLE MEAN

A frequency distribution of sample means is as follows.

$\bar{X}$	f
2	1
3	2
4	3
5	4
6	3
7	2
8	1



The mean of the sample means, denoted by

$$\mu_{\overline{x}} = \frac{2+3+\cdots+8}{16} = \frac{80}{16} = 5$$

which is the same as the population mean. Hencece,

$$\mu_{\overline{x}} = \mu$$

The standard deviation of sample means, denoted by $\sigma_{\overline{X}}$, is

$$\sigma_{\overline{X}} = \frac{\sqrt{(2-5)^2 + (3-5)^2 + \cdots + (8-5)^2}}{16} \approx 1.581$$

which is the same as the population standard deviation, divided by $\sqrt{2}$:

$$\sigma_{\overline{X}} = \frac{2.236}{\sqrt{2}} \approx 1.581$$

THE CENTRAL LIMIT THEOREM

As the sample size n increases, the shape of the distribution of the sample means taken with replacement from a population with mean μ and standard deviation σ will approach a normal distribution.

The mean of the sample means equals the population mean.

$$\mu_{\bar{X}} = \mu$$

The standard deviation of the sample means is called the **standard error of the mean**.

$$\sigma_{\bar{X}} = \sigma / \sqrt{n}.$$

A third property of the sampling distribution of sample means pertains to the shape of the distribution and is explained by the central limit theorem.

Used to gain information about an individual data value when the variable is normally distributed

$$z = \frac{x - \mu}{\sigma}$$

Used to gain information when applying the central limit theorem about a sample or more

$$z = \frac{\bar{X} - \mu}{\frac{\sigma}{\sqrt{n}}}$$

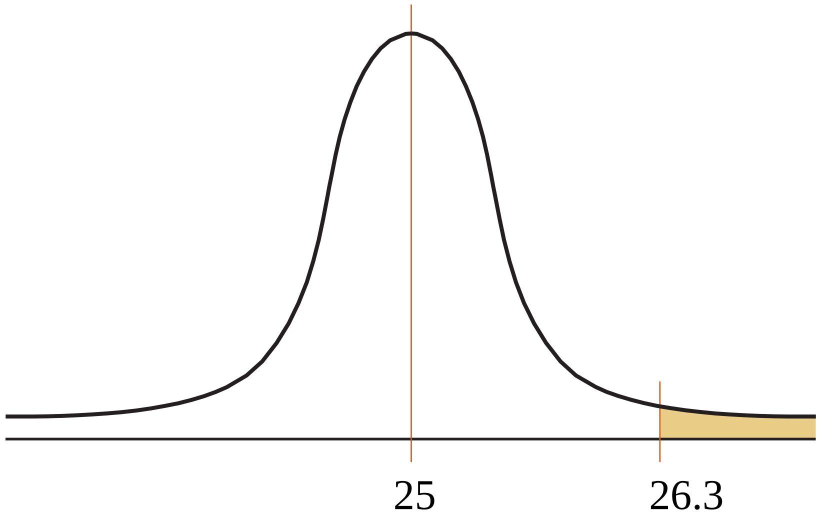
EXAMPLE 6–13HOURS THAT CHILDREN WATCH TELEVISION

A. C. Nielsen reported that children between the ages of 2 and 5 watch an average of 25 hours of television per week. Assume the variable is normally distributed and the standard deviation is 3 hours. If 20 children between the ages of 2 and 5 are randomly selected, find the probability that the mean of the number of hours they watch television will be greater than 26.3 hours.

SOLUTION

$\mu = 25$, $\sigma = 3$, n = 20

$$\sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}} = \frac{3}{\sqrt{20}} = 0.671$$



$$P\,(\bar{X} > 26.3) = 1 - P\left(\frac{\bar{X} - \mu}{\sigma/\sqrt{n}} < \frac{26.3 - 25}{3/\sqrt{20}}\right)$$

$$= 1 - \text{P}\,(z < 1.94)$$

$$= 1 - 0.9738$$

$=$ 0.0262 or 2.62%

The probability that the 20 children selected between the ages of 2 and 5 watch more than 26.3 hours of television per week is 2.62%.

EXAMPLE 6–14DRIVE TIMES

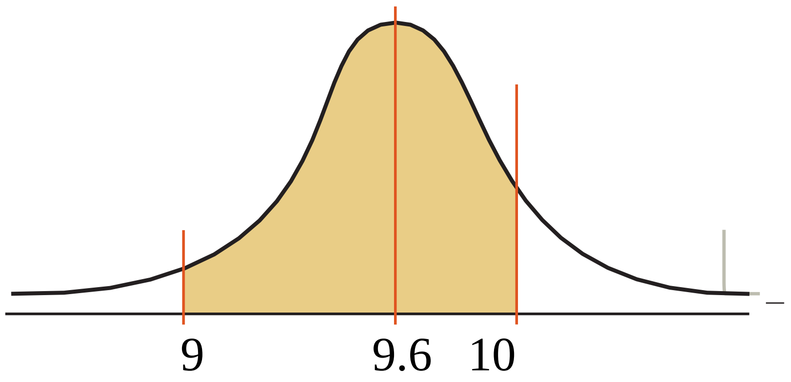
The average drive to work is 9.6 miles. Assume the standard deviation is 1.8 miles. If a random sample of 36 employed people who drive to work are selected, find the probability that the mean of the sample miles driven to work is between 9 and 10 miles.

x

Solution

$\mu = 9.6$, $\sigma = 1.8$, n = 36

$$\sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}} = \frac{1.8}{\sqrt{36}} = 0.3$$



$$P\ (9 < \bar{X} < 10) \qquad = \quad P\left(\frac{9-9.6}{1.8/\sqrt{36}} < \frac{\bar{X}-\mu}{\sigma/\sqrt{n}} < \frac{10-9.6}{1.8/\sqrt{36}}\right)$$

$$= \ \mathrm{P}\ (-2 < \mathrm{z} < 1.33)$$

= P(z< 1.33) – P(z < -2)

= 0.9082 – 0.0228

= 0.8854 or 88.54%

The probability that the mean mileage driven to work for a sample size of 36 is between 9 and 10 miles is 88.54%.

EXAMPLE 6–15 WORKING WEEKENDS

The average time spent by construction workers who work on weekends is 7.93 hours (over 2 days). Assume the distribution is approximately normal and has a standard deviation of 0.8 hour.

a. Find the probability that an individual who works at that trade works fewer than 8 hours on the weekend.

b. If a sample of 40 construction workers is randomly selected, find the probability that the mean of the sample will be less than 8 hours.

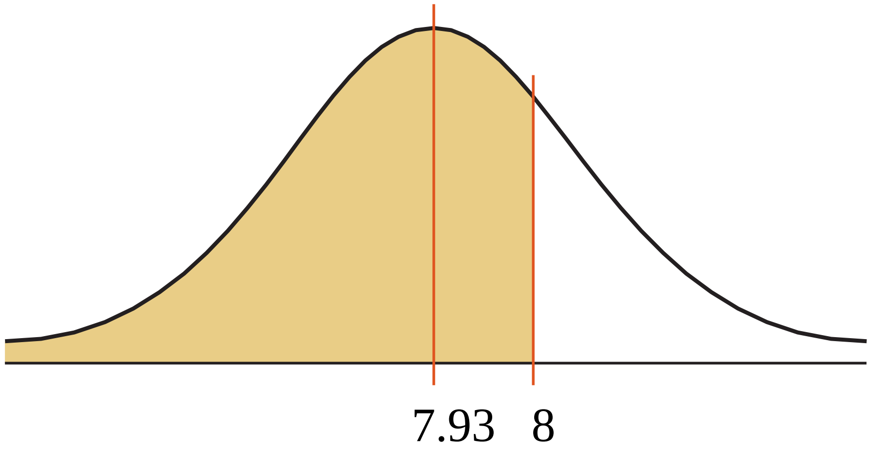
Solution

$\mu = 7.93$, $\sigma = 0.8$, n = 40

$$P(x < 8) = P\left(\frac{x-\mu}{\sigma} < \frac{8-7.93}{0.8}\right)$$

$= P(z < 0.09)$

$= 0.5359$ or 53.59%

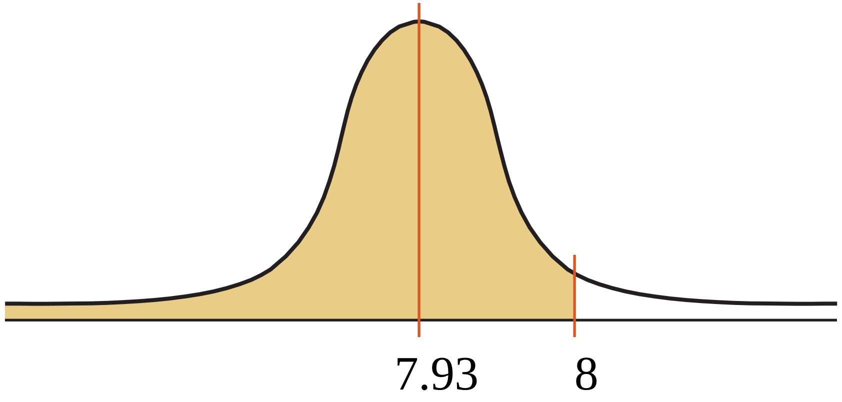


Hence, the probability of selecting a construction worker who works less than 8 hours on a weekend is 0.5359, or 53.59%.

$$P(\bar{X} < 8) = P\left(\frac{\bar{X}-\mu}{\sigma/\sqrt{n}} < \frac{8-7.93}{0.8/\sqrt{40}}\right)$$

$= P(z < 0.55)$

$= 0.7088$ or 70.88%



Hence, the probability of getting a sample mean of less than 8 hours when the sample size is 40 is 0.7088, or 70.88%.


Good Night


Thanks